RATIONAL METHODS IN MATRIX EQUATIONS[1]

MARK H. INGRAHAM

I. **Introduction.**[2] I shall start with what I hope will not prove an overelaborate statement of the limitations of this paper in scope and treatment. I shall assume throughout a knowledge of the definitions of a field, a division algebra, of matrices and rational operations thereon. I shall also need to assume a knowledge of what is meant by the invariant factors and elementary divisors of a matrix. Little essential will be lost if the only field considered by the listeners is the rational number system and the only division algebra that of quaternions over the rational field.

Consider a field $\mathfrak{K}$ and a system of constant matrices $A_1, \cdots, A_l$, unknown matrices $X_1, \cdots, X_n$ and equations

$$\phi_i(A_1, \cdots, A_l, X_1, \cdots, X_n) = 0 \tag{1}$$

where the ϕ_i's are polynomials with coefficients in $\mathfrak{K}$. If the elements of A_i are a_{ijk} and of X_i are x_{ijk}, (1) is equivalent to a system

$$\psi_s(\cdots, x_{ijk}, \cdots) = 0 \tag{2}$$

where the ψ_s are polynomials with coefficients in $\mathfrak{K}$. (If $\mathfrak{K}$ is replaced by a division algebra $\mathfrak{d}$ over $\mathfrak{K}$, the number of equations in (2) is merely enlarged.) We have therefore "reduced" the equations (1) to those of (2). This process we shall technically designate without great exaggeration as the "worst possible algorithm," or, following modern style, W.P.A. This indicates that no "tour de force" which shows that ultimately a matrix problem can be solved in a finite time, but shows little else, is of interest. This is a topic in which the above simple proof of the existence of inelegant methods means that we need only pay attention to results that essentially use the matricial properties of matrices, only to results and methods having at least a minimum degree of elegance.

Partly as a consequence of the above, this lecture, as is often the case, is not so much a description of broad theories of the nature we desire, as a report on what special cases have been found to be seduci-

[1] An address delivered before the Chicago meeting of the Society, April 14, 1939, by invitation of the Program Committee; presented in part April 15, 1939, under the title *An algorithm for the solution of the unilateral matrix equation.*

[2] The author wishes to thank H. C. Trimble, C. J. Everett, and J. H. Bell for help in preparing this paper. Their aid was made possible by the Research Committee of the University of Wisconsin.

ble. As a definition of the type of result I shall describe and the literary style in which this description will be given, I will lay down, for this lecture, four canons. The first and last of these I would not defend as completely valid dicta.

1. Irrational methods will be relegated to the background. In particular, we will not assume that we can reduce every polynomial into linear factors.

2. A theorem that is valid for the field of rational numbers is ipso facto of interest. This is particularly true if it extends to the case of all division algebras with finite basis over that field. (If a "field" had been defined to have characteristic zero and others to be "semi-fields," we would, I believe, have had a more wholesome set of values than at present.)

3. Proofs that are essentially algorithmic belong to the aristocracy.

4. The language of matrices, bases, vectors, linear spaces, polynomials, greatest common divisors, and so on, will be used instead of operators, modules, ideals, lattices, and so on. This is in deference to (a) the non-algebraists, (b) the frequency with which generalizations avoided are merely formal, (c) my personal taste.

The paper when presented as a symposium address included the discussion of the equation $TA=BT+C$ and a discussion of the unilateral matrix equation. Only the latter is included here since an extended form of the former is being published elsewhere. (Ingraham [8].)

II. **The unilateral equation.** We will now pass to the consideration of the unilateral (as to position of coefficients) equation

$$\sum A_i X^i = 0. \tag{3}$$

We will endeavor to show that much more can be said of this equation than has heretofore been shown. However, at best the general treatment is not simple so that it is desirable to use when available simpler theories for special types of equations, or for yielding special types of solutions.

What follows will be divided into three parts:

1. The case where all the A_i are polynomials in a single matrix A with scalar coefficients, and solutions X are sought which are of the same type.

2. The case where (3) is of the type

$$\psi(X) = A \tag{4}$$

where the polynomial ψ has scalar coefficients.

3. The general unilateral equation.

The solution of case 1 was first given by Szücs[3] and though his methods are not rational, the rewriting in rational form, as here done, is not difficult. This type of solution had been given, however, for (4) by Roth,[4] and later by a method similar to that of Szücs by Franklin.[5]

Consider

$$\phi(A, X) = 0 \tag{5}$$

and consider only such solutions X as are polynomials in A. We will also assume that the field $\mathfrak{F}$ has characteristic zero. Let $X(\lambda)$ be a polynomial such that $X(A)$ is a solution of (5). Hence $\phi(A, X(A)) = 0$. A necessary and sufficient condition that this be true is that $\phi(\lambda, X(\lambda))$ be divisible by the first invariant factor of A, that is, the minimal polynomial h such that $h(A) = 0$. Let g^k be the highest power of an irreducible polynomial g contained in h as a factor. Hence if $\phi(A, X(A)) = 0$, g^k must divide $\phi(\lambda, X(\lambda))$. Since $\mathfrak{F}$ has characteristic zero this is equivalent to g dividing $\phi, \phi', \phi'', \cdots, \phi^{(k-1)}$. (In case the characteristic is not zero the condition of divisibility is still workable but not as neat.)

Let $X(\lambda) = X_0(\lambda) + X_1(\lambda)g + X_2(\lambda)g^2 + \cdots$ where the X_i are reduced mod g.

Hence

$$\phi(\lambda, X_0) \equiv 0 \bmod g,$$

or in other words, $\phi(\lambda, X)$ has a zero in the field $\mathfrak{F}[\lambda]/\{g(\lambda)\}$ for every irreducible factor g of h.

Consider

$$\begin{aligned}\phi'(\lambda, X) &= \phi_\lambda(\lambda, X) + \phi_X(\lambda, X)X' \\ &\equiv \phi_\lambda(\lambda, X_0) + \phi_X(\lambda, X_0)(X_0' + X_1 g') \\ &\equiv (X_1 g')\phi_X(\lambda, X_0) + \psi_1 \bmod g\end{aligned}$$

where ψ_1 is fixed by the determination of X_0 and, in general,

$$\phi^{(s)}(\lambda, X) \equiv s!X_s(g')^s\phi_X(\lambda, X_0) + \psi_s \bmod g$$

where ψ_s is determined in terms of $X_0, \cdots, X_{s-1}$.

Note that since g is irreducible, g' is prime to g.

It is therefore necessary if $k > 1$ that either

$$\phi_X(\lambda, X_0) \not\equiv 0 \bmod g \tag{6}$$

[3] Szücs [7].

[4] Roth [6].

[5] Franklin [1].

or that

$$\psi_i \equiv 0 \bmod g, \qquad i = 1, \cdots, k-1. \tag{7}$$

If for every irreducible factor g these conditions are satisfied, then we have for each g_i an X_{g_i} such that X is a solution of $\phi(A, X) = 0$ if and only if

$$X \equiv X_{g_i} \bmod g_i^{k_i}.$$

That such an X exists follows from the Chinese Remainder Theorem for polynomials. Hence

THEOREM 1. *If $\mathfrak{R}$ has characteristic 0, the equation $\phi(A, X) = 0$ has a solution which is a polynomial in A if and only if for every irreducible factor g of the minimal polynomial of A there exist in $\mathfrak{R}[\lambda]$ solutions of $\phi(\lambda, X) \equiv 0 \bmod g$ which, in case g is a multiple factor of the minimal polynomial of A, satisfy* (6), *or* (7).

If the equation considered is of the form $\theta(X) = A$ where θ has scalar coefficients, then $\phi'(\lambda, X_0) = -1$ and (6) above becomes necessary so that we can state

THEOREM 2. *If $\mathfrak{R}$ has characteristic 0, the equation $\theta(X) = A$ has a solution which is a polynomial in A if and only if $\theta(X) \equiv \lambda \bmod g_i$ has a solution in $\mathfrak{R}[\lambda]$, for every irreducible factor g_i of the minimum function of A which, in case g_i is a multiple factor of the minimal function of A, does not satisfy $\theta'(X) \equiv 0 \bmod g_i$.*

As a corollary of this,

THEOREM 3. *If $\mathfrak{R}$ is algebraically closed and has characteristic 0, the equation $X^n = A$ has a solution $X(A)$, a polynomial in A, if and only if 0 is not a multiple root of the minimum function of A.*

It is not difficult to compute the number of such functions.

For certain cases the existence of solutions not polynomials in A is easily established by example.

For instance

$$X^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

has the solution

$$X = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

It is interesting, therefore, to study the general solutions of equations of the type discussed above. This is readily done for the equations $\theta(X)=A$, if A is nonderogatory, that is, its minimum equation is its characteristic equation. Since A is nonderogatory and every solution X is commutative with A, X is a polynomial in A. In this case our problem is completely solved. In case ϕ has scalar coefficients, the general solution of

$$\phi(X) = A$$

is determined as follows.[6]

If we find a matrix Y such that $\phi(Y)$ is similar to A, then a nonsingular matrix S may be found such that $A=S\phi(Y)S^{-1}=\phi(SYS^{-1})$ and hence SYS^{-1} is effective as X. Consider the invariant factors of $A-\lambda I$, $h_1, h_2, \cdots, h_k$. Then $\phi(Y)$ is similar to A if and only if the nullity of $h_i(\phi(Y))$ equals the nullity of $h_i(A)$. Hence the invariant factors of Y must be divisors of $h_1(\phi(\lambda))$. Under this restriction and the fact that the nullity of any polynomial ϕ in Y is determined as the sum of the degrees of the greatest common divisors of the invariant factors of Y, with the polynomial ϕ, diophantine equations may be written, any solution of which determines the invariant factors of a solution X, and hence determines X to within a transformation by a matrix commutative with A. All solutions are the transforms by matrices commutative with A of a certain finite dissimilar system of solutions $X_1, X_2, \cdots, X_k$. Those X_i which are polynomials in A are those solutions which are dissimilar to all other solutions. For example if

$$X^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

then

$$X = \begin{pmatrix} 0 & 0 & 0 \\ m & 0 & k \\ 1/k & 0 & 0 \end{pmatrix}$$

with $k\neq 0$.

This work generalizes with only moderate difficulty to the case where the scalars involved belong to a division algebra provided we ask the question correctly.[7] In the case where the scalars are in a field and $\phi=\sum\lambda^i a_i$ then $\phi(X)=A$ is equivalent to $\phi(X)\xi=\sum X^i\xi a_i$

[6] Ingraham [2].
[7] Ingraham [3].

$=A\xi$ for every vector ξ. This is not the case if the a_i belong to a noncommutative division algebra $\mathfrak{D}$. The correct question which can be answered is not to find a solution of $\phi(X)=A$, but given a set $[a_i]$ in $\mathfrak{D}$ to find a matrix X such that $\sum X^i\xi a_i=A\xi$ for every ξ. This problem turns out to be equivalent to a system of equations

$$\phi_i(X) = A_i$$

where the coefficients of ϕ_i are in the centrum of the division algebra $\mathfrak{D}$.

Let us now consider the general unilateral matrix equation

$$\sum A_iX^i = 0, \tag{8}$$

where the A_i are $n\times n$ matrices of constants.

Fundamental to the consideration of (8) is the fact that if X is any constant matrix then $A(\lambda)=\sum A_i\lambda^i$ can be expressed in one and only one way as

$$(\sum B_i\lambda^{i-1})(\lambda I - X) + B_0$$

and in fact

$$B_0 = \sum A_iX^i.$$

Hence

THEOREM 4. *X is a solution of $\sum A_iX^i=0$ if and only if $\lambda I-X$ is a right factor of the matrix $A(\lambda)=\sum A_i\lambda^i$.*

Theorem 5, which is a corollary of a theorem due to Phillips,[8] yields us an algorithm better than the W.P.A. This theorem is as follows:

THEOREM 5. *If X is a solution of $\sum A_iX^i=0$, then the determinant of the matrix $\lambda I-X$ divides the determinant of the matrix $\sum A_i\lambda^i$.*

This is, of course, a corollary of Theorem 4. The proof originally given was far more complicated than this proof which is due to C. C. MacDuffee. Phillips' theorem yields a necessary condition on the invariant factors of any solution. If Y has such invariant factors and $X=PYP^{-1}$ is a solution, then $\sum A_iPY^i=0$. The problem is then reduced to finding all P's satisfying this equation for canonical Y's with possible invariant factors.

Far stronger conditions can be secured from Theorem 6, which is a consequence of the following lemma.

[8] Phillips [5].

LEMMA. *If A, B, C, are λ-matrices such that $A=BC$, then the invariant factors of B (or C) divide the corresponding invariant factors of A.*

Thus

THEOREM 6. *If X is a solution of $\sum A_i X^i=0$, then the invariant factors of the matrix $\lambda I-X$ are divisors of the corresponding invariant factors of the matrix $\sum A_i\lambda^i$.*

This eliminates many cases not strained out by the sieve of Phillips' theorem.

That these conditions or any conditions in terms of the invariant factors of $A(\lambda)$ are insufficient to guarantee the existence of solutions with a given set of invariant factors for $\lambda I-X$ is shown by the two equations

$$\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}X^2+\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}=0$$

and

$$X+\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}=0.$$

The λ-matrices for these are respectively

$$\begin{pmatrix}\lambda^2 & 0\\ 0 & 1\end{pmatrix},\quad \begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix},$$

each of which has invariant factors λ^2, 1; but the first equation has no solution and the second has the solution

$$\begin{pmatrix}0 & -1\\ 0 & 0\end{pmatrix}.$$

We now pass to the description of an algorithm believed to be new for the solution of (8).

We shall call a matrix with elements polynomials in λ unimodular if its determinant is a scalar not zero, or what is equivalent, if it is nonsingular in the field of rational functions of λ, and if its reciprocal has elements polynomials in λ.

We say that A is in triangular form if all the elements below the main diagonal are zero. We say it is in canonical triangular form if it is in triangular form and all the elements above the main diagonal are of lower degree than the elements in the same column on the main diagonal and if when a zero occurs on the main diagonal the whole row in which it occurs is zero. We also specify that the leading coefficients of the polynomials along the main diagonal be 1.

If U is unimodular, then $\lambda I - X$ is a right factor of A if and only if it is a right factor of UA. Moreover, a unimodular matrix U may be found such that UA is in canonical triangular form, this form being completely determined by[9] A. From now on we will assume that A is in such form.

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \cdots a_{1n} \\ 0 & a_{22} & a_{23} \cdots a_{2n} \\ 0 & 0 & a_{33} \cdots a_{3n} \\ \cdot & \cdot & \cdot \quad \cdots \cdot \\ \cdot & \cdot & \cdot \quad \cdots \cdot \\ 0 & 0 & 0 \quad \cdots a_{nn} \end{pmatrix}$$

where the a_{ij}'s are polynomials in λ.

If $A = H(\lambda I - X)$ and if U is the unimodular matrix such that $U(\lambda I - X)$ is in canonical triangular form, then

$$A = HU^{-1}U(\lambda I - X)$$

and $T = U(\lambda I - X)$ is a right factor of A.

Our problem is therefore reduced to finding the triangular factors of A which are the canonical triangular forms of matrices of type $\lambda I - X$ where X is independent of λ.

We study then the problem of finding canonical triangular factors T of the matrix A and then of selecting those which are the canonical triangular forms of matrices of type $\lambda I - X$ for some X.

If $A = ST$ where the matrices A and T are in triangular form, S is in triangular form, and

$$(9) \qquad a_{i,i+j} = \sum_{k=0}^{j} s_{i,i+k} t_{i+k,i+j}, \qquad i = 1, \cdots, n; j = 0, \cdots, n - i.$$

For the case $j = 0$ these relations reduce to

$$a_{ii} = s_{ii} t_{ii}$$

and hence to the fact that the diagonal elements of T are factors of the corresponding diagonal elements of A.

Equations (9) give linear congruences each conditioning (not necessarily uniquely) the possible t_{ij} in terms of previously determined elements of T below t_{ij} occurring in the jth column and elements of s to the left of s_{ij} in the ith row—these may be considered therefore

[9] MacDuffee [4].

as recursive formulae for the t_{ij}. In these congruences the coefficient of t_{ij} is s_{ii} and the modulus is t_{jj}, so that, in case $\mathfrak{R}$ has infinitely many elements, an infinitude of solutions exists if and only if s_{ii} and t_{jj} are not relatively prime and at least one solution exists.

It should be noted that if T is the triangular form of M, the product of the first i diagonal terms is the greatest common divisor of the determinants of the ith order minors of the first i columns of M. Hence if T is of the form $U(\lambda I - X)$ where U is unimodular, the sum of the degrees of the first i terms of the principal diagonal of T must be less than or equal to i and the sum of the degrees of all the diagonal terms is n, the order of T. I believe in most equations written at random these last conditions would preclude a solution.

We turn now to determining X such that a T satisfying the above conditions is the triangular canonical form for some $(\lambda I - X)$. It can be shown (but the proof is omitted) that X is unique if it exists.

If $T = U^{-1}(\lambda I - X)$ where U is unimodular, then

$$U = (\lambda I - X)T^{-1}. \tag{10}$$

If we choose a matrix X, $(\lambda I - X)T^{-1}$ will have elements rational functions of λ but not, in general, polynomials.

The problem therefore reduces to seeking X such that $(\lambda I - X)T^{-1}$ is a matrix with elements polynomials in λ.

Let $T^{-1} = F/d$ where F is in triangular form and d is the determinant of T. The f_{ij} will be polynomials.

If (10) is fulfilled

$$u_{ij} = (\lambda f_{ij} - \sum x_{ik} f_{kj})/d;$$

hence u_{ij} is a polynomial if and only if

$$\sum x_{ik} f_{kj} \equiv \lambda f_{ij} \bmod d.$$

Since the f_{ij} are of degree less than d, these linear congruences are easily reduced, by equating coefficients, to n systems of linear equations for the x_{ik} (i fixed) all of the systems having the same matrix of coefficients, the rank of which must (in light of the uniqueness theorem) be n if a solution of all the systems exists.

Since this paper was delivered at Chicago Mr. J. H. Bell has furnished me with the following note on the conditions upon T for a matrix X to exist such that T is a triangular form of $\lambda I - X$.

The existence of X depends directly upon the coefficients of the elements of T.

We may obtain an $n \times n$ matrix D whose elements are scalars by the following steps.

1. Augment each element of T by the proper powers of λ with zero coefficients so that terms of the same degree as in the corresponding diagonal element appear.

2. Break up each column into separate columns, each one of which involves only monomials of the same degree in λ, for example,

$$\begin{pmatrix} a\lambda + b \\ c\lambda + d \end{pmatrix} \rightarrow \begin{pmatrix} a\lambda, & b \\ c\lambda, & d \end{pmatrix}.$$

3. Delete the columns which do not involve λ.

4. Set $\lambda = 1$ obtaining D.

A necessary and sufficient condition that there exist a matrix X such that $T = U^{-1}(\lambda I - X)$ is that $|D| \neq 0$.

If $T = \sum_r^0 T_i \lambda^i$ ($T_i^{n \times n}$ involving scalars only), then the above condition is equivalent to saying that n is the rank of

$$M = (T_r, T_{r-1}, \cdots, T_2, T_1).$$

X may be obtained quite readily from D. We first find $B = (b_{ik})$ as follows. If $t_{jj} \neq 1$, $(b_{j1} \cdots b_{jn}) = (0 \cdots 0, -1, 0 \cdots 0)D^{-1}$ where the -1 occurs in the position corresponding to the column of D containing $t_{1,jj}$ (where $t_{kj} = \sum_i^0 t_{i,kj}\lambda^i$). If $t_{ll} = 1$, $(b_{l1} \cdots b_{ln}) = \eta_l D^{-1}$ where η_l is a $1 \times n$ vector obtained by taking the lth row of the matrix obtained from the matrix in step 2 by deleting the columns involving the leading term of t_{kk}, for every k, and setting $\lambda = 1$.

Then

$$X = BT_0.$$

Bibliography

1. P. Franklin, *Algebraic matric equations*, Journal of Mathematics and Physics, Massachusetts Institute of Technology, vol. 10 (1932), pp. 289–314.

2. M. H. Ingraham, *On the rational solution of the matrix equation* $P(X) = A$, Journal of Mathematics and Physics, vol. 13 (1934), pp. 46–50.

3. M. H. Ingraham, *On certain equations in matrices whose elements belong to a division algebra*, this Bulletin, vol. 44 (1938), pp. 117–124.

4. C. C. MacDuffee, *Matrices with elements in a principal ideal ring*, this Bulletin, vol. 39 (1933), pp. 564–584.

5. H. B. Phillips, *Functions of matrices*, American Journal of Mathematics, vol. 41 (1919), pp. 266–278.

6. W. E. Roth, *A solution of the matric equation* $P(X) = A$, Transactions of this Society, vol. 30 (1928), pp. 579–596.

7. A. Szücs, *Sur les équations définissant une matrice en fonction algébrique d'une autre*, Acta Litterarum ac Scientiarum, Szeged, vol. 7 (1934–1935), pp. 48–50.

8. M. H. Ingraham and H. C. Trimble, *On the matric equation* $TA = BT + C$, American Journal of Mathematics, vol. 63, pp. 9–27.

University of Wisconsin

MAXIMUM OF CERTAIN FUNDAMENTAL LAGRANGE INTERPOLATION POLYNOMIALS[1]

M. S. WEBSTER

This note extends some of the results obtained in a previous paper[2] which we shall designate as I. The notations are the same.

We are concerned with the polynomials

$$l_k^{(n)}(x) \equiv \frac{\phi_n(x)}{\phi_n'(x_k)(x - x_k)}, \qquad k = 1, 2, \cdots, n,$$

where $\phi_n(x) \equiv (x-x_1)(x-x_2)\cdots(x-x_n)$ is the Jacobi polynomial of degree n which satisfies the differential equation $(1-x^2)\phi_n''(x) + [\alpha-\beta-(\alpha+\beta)x]\phi_n'(x) + n(n+\alpha+\beta-1)\phi_n(x) = 0$. The parameters α, β are positive and n is a positive integer. It is known that $-1 < x_n < x_{n-1} < \cdots < x_1 < 1$. Throughout the paper, x is always restricted to the interval $-1 \leqq x \leqq 1$.

It was shown in I, for example, that, if $\alpha = \beta = \frac{3}{2}$, $\max |l_k^{(n)}(x)| < 2$ and $l_1^{(n)}(1) \to 2$ as $n \to \infty$.

Now we use[3]

$$\phi_n(1) = \frac{2^n\Gamma(n+\beta)\Gamma(n+\alpha+\beta-1)}{\Gamma(\beta)\Gamma(2n+\alpha+\beta-1)}$$

and the asymptotic expressions[4]

$$\phi_n(\cos\theta) = \frac{2^n\Gamma(n+1)\Gamma(n+\alpha+\beta-1)}{(\pi n)^{1/2}\Gamma(2n+\alpha+\beta-1)}\left(\sin\frac{\theta}{2}\right)^{1/2-\beta}\left(\cos\frac{\theta}{2}\right)^{1/2-\alpha} \cdot \left\{\cos\left[N\theta - (2\beta-1)\pi/4\right] + (n\sin\theta)^{-1}O(1)\right\},$$

$$\phi_n(\cos\theta) = \frac{2^n\Gamma(n+1)\Gamma(n+\alpha+\beta-1)}{\Gamma(2n+\alpha+\beta-1)}\left(\sin\frac{\theta}{2}\right)^{1-\beta}\left(\cos\frac{\theta}{2}\right)^{1-\alpha} \left\{\frac{\Gamma(n+\beta)}{\Gamma(n+1)}\left(\frac{\theta}{\sin\theta}\right)^{1/2}\frac{J_{\beta-1}(N\theta)}{N^{\beta-1}} + \theta^{1/2}O(n^{-3/2})\right\},$$

where $N = n + (\alpha+\beta-1)/2$, $cn^{-1} \leqq \theta \leqq \pi - \epsilon$, c, ϵ positive constants and

[1] Presented to the Society, April 13, 1940.

[2] M. Webster, *Note on certain Lagrange interpolation polynomials*, this Bulletin, vol. 45 (1939), pp. 870–873.

[3] C. Winston, *On mechanical quadratures formulae involving the classical orthogonal polynomials*, Annals of Mathematics, (2), vol. 35 (1934), pp. 658–677.

[4] G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. 23, 1939, pp. 191–192, 121, 123.

where $J_m(x)$ is Bessel's function of order m. Since $\phi_n'(x;\alpha,\beta) = n\phi_{n-1}(x;\alpha+1,\beta+1)$, these yield immediately the following results:

LEMMA. *For x_k such that $-1+\epsilon \leqq x_k \leqq 1-\epsilon$ and $|x-x_k| \geqq \epsilon' > 0$, $\max |l_k^{(n)}(x)| \to 0$ as $n\to\infty$ even if $x\to\pm 1$ ($\alpha,\beta<\frac{3}{2}$; $\epsilon,\epsilon'>0$).*

THEOREM 1. *For x_k such that $-1+\epsilon \leqq x_k \leqq 1-\epsilon$ and $|x-x_k| \geqq \epsilon' > 0$, $\max |l_k^{(n)}(x)| = O(n^\gamma)$ as $n\to\infty$ where $\max(\alpha,\beta) = \gamma > \frac{3}{2}$; $\epsilon,\epsilon'>0$. The exponent γ cannot be decreased.*

The method used in the proof of Theorem 5 in I really gives the following slightly stronger result:

THEOREM 2. *If $-1+\epsilon \leqq x_k \leqq 1-\epsilon$, $-1+\epsilon' \leqq x \leqq 1-\epsilon'$, $\max |l_k^{(n)}(x)| \to 1$ as $n\to\infty$ ($\epsilon,\epsilon'>0$).*

Combining Theorem 2 and the lemma, we obtain the following:

THEOREM 3. *For x_k such that $-1+\epsilon \leqq x_k \leqq 1-\epsilon$, $\max |l_k^{(n)}(x)| \to 1$ as $n\to\infty$ ($\alpha,\beta<\frac{3}{2}$, $\epsilon>0$).*

This result is a considerable improvement over Theorem 5 in I. Moreover, if the hypothesis $-1+\epsilon \leqq x_k \leqq 1-\epsilon$ is removed, the theorem is not true as Erdös and Grünwald[5] showed in case $\alpha=\beta=\frac{1}{2}$. In view of Theorems 1, 4, 5, 6, the restriction $\alpha,\beta<\frac{3}{2}$ is also necessary. In particular, this theorem holds for the case of Tschebycheff ($\alpha=\beta=\frac{1}{2}$) and Legendre ($\alpha=\beta=1$) polynomials.

THEOREM 4. *If $\alpha=\beta=\frac{3}{2}$ and $x_k\to t$ as $n\to\infty$, then $\max |l_k^{(n)}(x)| \to 1+|t|$ as $n\to\infty$ ($-1\leqq t\leqq 1$). This is also an upper bound if $|x_k| < |t|$ at least for large values of n.*

PROOF. It was shown in I that $l_k^{(n)}(1) = 1+x_k$ and for $x_{k+1} \leqq x \leqq x_{k-1}$, $\max |l_k^{(n)}(x)| < 1.87$. Since (I) $\max |l(_k^{(n)}x)|$ is attained either between x_{k+1} and x_{k-1} or at $x=\pm 1$, the theorem is valid for $t=1$ and, by symmetry, for $t=-1$.

If $|t|<1$, the preceding paragraph and Theorem 2 complete the proof. In fact, $\max |l_k^{(n)}(x)| = 1+|x_k|$ at least for large n.

The next two theorems are obtained in a similar manner.

THEOREM 5. *If $\alpha=\frac{1}{2}$, $\beta=\frac{3}{2}$ and $x_k\to t$ as $n\to\infty$, then $\max |l_k^{(n)}(x)| \to 4/\pi$ if $t=-1$, 1 if $-1<t\leqq -\frac{1}{2}$, $(2(1+t))^{1/2}$ if $-\frac{1}{2}\leqq t\leqq 1$.*

THEOREM 6. *If $\alpha=\frac{3}{2}$, $\beta=\frac{1}{2}$ and $x_k\to t$ as $n\to\infty$, then $\max |l_k^{(n)}(x)| \to [2(1-t)]^{1/2}$ if $-1\leqq t\leqq\frac{1}{2}$, 1 if $\frac{1}{2}\leqq t<1$, $4/\pi$ if $t=1$.*

[5] Erdös and Grünwald, *Note on an elementary problem of interpolation*, this Bulletin, vol. 44 (1938), pp. 515–518.

The max $|l_1^{(n)}(x)|$ is attained at $x=\pm 1$ since[4] (I) $\theta_{k+1}-\theta_k \leqq 2\pi/(2n+\alpha+\beta-1)$ provided $\frac{1}{2}\leqq\alpha$, $\beta\leqq\frac{3}{2}$ and $x_k\equiv\cos\theta_k$. Using the second asymptotic formula and the fact[4] that $n\theta_k\to j_k$ as $n\to\infty$ where j_k is the kth positive zero of $J_{\beta-1}(x)$, we find that

$$|l_k^{(n)}(1)| \to (\tfrac{1}{2}j_k)^{\beta-2}\,|\Gamma(\beta)J_\beta(j_k)|^{-1} \qquad \text{as } n\to\infty,\ k \text{ constant},$$

$l_1^{(n)}(-1)\to 0$ which proves the theorem:

THEOREM 7. *Max* $|l_1^{(n)}(x)| \to (\frac{1}{2}j_1)^{\beta-2}|\Gamma(\beta)J_\beta(j_1)|^{-1}$ *as* $n\to\infty$ (*where* $\frac{1}{2}\leqq\alpha$, $\beta\leqq\frac{3}{2}$, j_1 *is first positive zero of* $J_{\beta-1}(x)$).

A similar result holds for $l_n^{(n)}(x)$ if β is replaced by α.

For Legendre polynomials ($\alpha=\beta=1$) this limit is approximately 1.602. For $\alpha=\beta=\frac{1}{2}$ and $\alpha=\beta=\frac{3}{2}$ the limit of Theorem 7 is also an upper bound for max $|l_1^{(n)}(x)|$ and max $|l_k^{(n)}(x)|$. Whether this is true, in general, remains unanswered.

PURDUE UNIVERSITY

AN INVARIANCE THEOREM FOR SUBSETS OF S^n [1]

SAMUEL EILENBERG

The purpose of this paper is to establish the following.

INVARIANCE THEOREM. *Let A and B be two homeomorphic subsets of the n-sphere S^n. If the number of components of S^n-A is finite, then it is equal to the number of components of S^n-B.*

In the case when A and B are closed this theorem is a very well known consequence of Alexander's duality theorem and its generalizations. In our case we also derive our result as a consequence of a duality theorem. However, the duality is established only for the dimension $n-1$.

Given a metric space X we shall say that Γ^k is a k-cycle in X if there is a compact subset A of X such that Γ^k is a k-dimensional convergent (Vietoris) cycle in A with coefficients modulo 2. We shall write $\Gamma^k\sim 0$ if $\Gamma^k\sim 0$ holds in some compact subset of X. The homology group of X obtained this way will be denoted by $\mathfrak{H}^k(X)$; the corresponding connectivity number, by $p^k(X)$. The number $p^k(X)$ can be either finite or ∞.

[1] Presented to the Society, December 28, 1939.

DUALITY THEOREM. *Let $A \subset S^n$ and let $z_0, z_1, \cdots, z_m$ belong to $m+1$ different quasi-components*[2] *of $S^n - A$. There are m linearly independent (modulo 2) $(n-1)$-cycles*

$$\Gamma_1^{n-1}, \cdots, \Gamma_m^{n-1} \tag{1}$$

of A such that

$$v(\Gamma_i^{n-1}, \overset{0}{\gamma_j}) = \delta_{ij}, \qquad i, j = 1, \cdots, m, \tag{2}$$

where γ_j^0 is the 0-cycle $z_0 + z_j$ (consisting of the two points z_0, z_j each of them with coefficient 1) and $v(\Gamma, \gamma)$ is the linking number.

In case $S^n - A$ has only $m+1$ quasi-components, the cycles (1) form a basis for $\mathfrak{C}^{n-1}(A)$.

PROOF. In case A is closed the theorem turns out to be a particular case of the generalized Alexander duality theorem.[3] We shall prove our theorem for arbitrary sets A using the theorem for closed sets.

Since $z_0, z_1, \cdots, z_m$ belong to $m+1$ different quasi-components of $S^n - A$ there is a decomposition $S^n - A = A_0 + A_1 + \cdots + A_m$ such that $z_i \in A_i$ and $A_i \bar{A}_j + \bar{A}_i A_j = 0$ for $i \neq j$, $i, j = 0, 1, \cdots, m$. Let $B_0, B_1, \cdots, B_m$ be open disjoint sets such that $A_i \subset B_i$ for $i = 0, 1, \cdots, m$ and let $B = S^n - (B_0 + B_1 + \cdots + B_m)$. Clearly B is a closed subset of A and $z_0, z_1, \cdots, z_m$ belong to $m+1$ different quasi-components (equals components) of $S^n - B$.

Applying the duality theorem to the closed set B we obtain the cycles (1) satisfying (2). In order to prove that they determine linearly independent elements modulo 2 of $\mathfrak{C}^{n-1}(A)$ consider a cycle $\Gamma^{n-1} = a_1\Gamma_1^{n-1} + \cdots + a_m\Gamma_m^{n-1}$ where $a_j = 0, 1$. It follows from (2) that $v(\Gamma^{n-1}, \gamma_j^0) = a_j$. Therefore $\Gamma^{n-1} \sim 0$ in A implies $a_1 = \cdots = a_m = 0$.

Suppose now that $S^n - A$ consists of exactly $m+1$ quasi-components. It follows that the sets $A_0, A_1, \cdots, A_m$ are connected.

Let Γ^{n-1} be an $(n-1)$-cycle of A contained in some closed set $D \subset A$. Let E_i be the component of $S^n - (B+D)$ containing A_i $(i = 0, 1, \cdots, m)$ and let $E = S^n - (E_0 + E_1 + \cdots + E_m)$. It follows that (1°) E is a closed subset of A, (2°) $S^n - E$ consists of exactly $m+1$ quasi-components (equals components), (3°) the points

[2] Two points $x_1, x_2 \in X$ belong to the same quasi-component of X if there is no decomposition $X = A_1 + A_2$ such that $x_1 \in A_1$, $x_2 \in A_2$ and $A_1\bar{A}_2 + \bar{A}_1 A_2 = 0$. If the number of quasi-components of X is finite then every quasi-component is a component.

[3] Alexander, J. W., Transactions of this Society, vol. 23 (1922), pp. 333–349; Frankl, F., Sitzungsberichte der Wiener Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Klasse, 2A, vol. 136 (1927), pp. 689–699; Alexandroff, P., Annals of Mathematics, (2), vol. 30 (1928), p. 163.

$z_0, z_1, \cdots, z_m$ belong to different quasi-components of $S^n - E$, (4°) the cycles (1) and Γ^{n-1} are contained in E. According to the duality theorem for closed sets the cycles (1) form a basis for $\mathfrak{C}^{n-1}(E)$. This implies the existence of $a_1, a_2, \cdots, a_m$ $(a_i = 0, 1)$ such that

$$\Gamma^{n-1} \sim a_1\Gamma_1^{n-1} + \cdots + a_m\Gamma_m^{n-1} \text{ in } E.$$

This proves the theorem since $E \subset A$.

Given a metric space X let the number $b_0(X)$ be defined as follows:

$b_0(X) = 0$ if $X = 0$,
$b_0(X) = m$ if $X \neq 0$ and X has exactly $m+1$ components,
$b_0(X) = \infty$ if X has an infinity of components.

Clearly the value of $b_0(X)$ remains unchanged if we replace in its definition components by quasi-components. The duality theorem implies therefore the following:

(I) *For every subset A of S^n we have*

$$p^{n-1}(A) = b_0(S^n - A).$$

(II) *For every two homeomorphic subsets A and B of S^n we have*

$$b_0(S^n - A) = b_0(S^n - B).$$

The invariance theorem stated in the introduction follows directly from (II).

If X consists of an infinity of components, then instead of taking $b_0(X) = \infty$ we could define $b_0(X)$ to be the cardinal number corresponding to the class of all components of X. Similarly $p^k(X)$ could be redefined as a cardinal number. But with these new definitions (I) and (II) are no longer true.[4] In fact, let A be a subset of S^1 such that $S^1 - A$ is closed and enumerably infinite, and let B be a subset of S^1 such that $S^1 - B$ is perfect and non-dense. It is clear that A and B are homeomorphic, that $b_0(S^1 - A) = p^0(A) = p^0(B) = \aleph_0$, and that $b_0(S^1 - B) = 2^{\aleph_0}$.

UNIVERSITY OF MICHIGAN

[4] That (II) is no longer true was first pointed out to me by Dr. L. Zippin.

MEASURABILITY AND MODULARITY IN THE THEORY OF LATTICES[1]

M. F. SMILEY

In a previous note[2] a notion of measurability (with respect to a function μ) of elements of an arbitrary lattice was introduced. Our purpose there was to study closure properties of the subset of measurable elements. To do this it was convenient to assume that the lattice was modular. Results of V. Glivenko,[3] and of L. R. Wilcox and the author[4] indicate that the idea of measurability and that of modularity are intimately related. The purpose of this note is to exhibit a further relationship which does not depend on metric properties of the function μ.

In a lattice L we call the ordered pair (a, b) of elements of L *modular* and write $(a, b)M$ in case $(a_1+a)b=a_1+ab$ for every $a_1 \leqq b$. This relation has been studied by L. R. Wilcox.[5] If $\mu(a)$ is a real valued function defined on L we say that an element $c \in L$ is μ-measurable in case

$$\mu(c) + \mu(b) = \mu(c + b) + \mu(cb)$$

for every $b \in L$. The symbol $L(\mu)$ will denote the totality of elements of L which are μ-measurable. We call μ *proper* in case $a \leqq b$ with $\mu(a)=\mu(b)$ implies $a=b$.

Before we discuss the general case, let us consider a lattice L_0 of finite dimension (that is, satisfying both chain conditions) in which every principal chain joining two elements a, $b \in L_0$ has the same length $n(a, b)$. Examples of such lattices are well known.[6] Let 0 de-

[1] Presented to the Society, February 24, 1940.

[2] *A note on measure functions in a lattice*, this Bulletin, vol. 46 (1940), pp. 239–241. This concept specializes to that of permutability with all subgroups (that is, the *quasi-normality* of O. Ore, *Structures and group theory* I, Duke Mathematical Journal, vol. 3 (1937), p. 162) if we set $\mu(s)=\log o(s)$ in the lattice of subgroups of a finite group, and to that of *measurability* in the sense of Carathéodory with respect to an *outer measure* function.

[3] *Contributions à l'étude des systèmes de choses normées*, American Journal of Mathematics, vol. 59 (1937), pp. 933–934.

[4] *Metric lattices*, Annals of Mathematics, (2), vol. 40 (1939), p. 313.

[5] *Modularity in the theory of lattices*, Annals of Mathematics, (2), vol. 40 (1939), p. 491 ff.

[6] The "exchange" lattices discussed by Saunders Mac Lane (*A lattice formulation for transcendence degrees and p-bases*, Duke Mathematical Journl, vol. 4 (1938), pp. 455–468) as well as the "semi-modular" lattices of finite dimension of Wilcox (op. cit., pp. 502–505) have these properties. We make no use, however, of the metric properties of these systems.

note the least element of L_0 and define $\mu_0(a) \equiv n(0, a)$. We note that μ_0 is proper. These considerations lead to the following theorem.

THEOREM 1. *If $c \in L_0$ then $c_1 \in L_0(\mu_0)$ for every $c_1 \leqq c$ if and only if $a \leqq c$ implies that $(b, a)M$ for every $b \in L_0$ and furthermore that $x = ax + b$ for every $x \in L_0$ for which $b \leqq x \leqq a + b$.*

PROOF. To show the sufficiency of the condition consider elements $b, c_1 \in L_0$ for which $c_1 \leqq c$. It is easily seen that $c_1 \in L_0(\mu_0)$ provided that $n(c_1 b, c_1) = n(b, c_1 + b)$. We shall prove in fact that $c_1/c_1 b \simeq c_1 + b/b$. As usual, for $x_1 \in c_1/c_1 b$, that is, $c_1 b \leqq x_1 \leqq c_1$, define $f(x_1) \equiv x_1 + b$. Clearly $f(x_1) \in c_1 + b/b$. Also $f(c_1/c_1 b) = c_1 + b/b$; for if $b \leqq x \leqq c_1 + b$, by hypothesis $f(c_1 x) = x$. Suppose $f(x_1) = f(x_2)$, that is, $x_1 + b = x_2 + b$. Clearly $(x_1 + b)c_1 = (x_2 + b)c_1$. By hypothesis $(b, c_1)M$, and it follows that $x_1 + bc_1 = x_2 + bc_1$, $x_1 = x_2$. Thus $f(x_1)$ establishes a one-to-one correspondence between $c_1/c_1 b$ and $c_1 + b/b$. This correspondence preserves the ordering relation and hence $c_1/c_1 b \simeq c_1 + b/b$. The proof of the sufficiency of the condition is complete. The proof of the necessity is included in the proof of the following theorem.

THEOREM 2. *Suppose that $\mu(a)$ is a proper real valued function defined on a lattice L. Consider an element $c \in L$. If $c_1 \in L(\mu)$ for every $c_1 \leqq c$, then $(b, a)M$ for every $a \leqq c$ and every $b \in L$, and furthermore $x = ax + b$ for every $x \in L$ such that $b \leqq x \leqq a + b$.*

PROOF. Consider elements $a, b \in L$ with $a \leqq c$. To show that $(b, a)M$, it suffices, since μ is proper, to prove that

$$\mu((a_1 + b)a) = \mu(a_1 + ab)$$

for every $a_1 \leqq a$. By hypothesis $a_1, a \in L(\mu)$ and consequently

$$\begin{aligned}
\mu((a_1 + b)a) &= \mu(a_1 + b) + \mu(a) - \mu(a + b) \\
&= \mu(a_1 + b) + \mu(ab) - \mu(b) \\
&= \mu(a_1) + \mu(b) - \mu(a_1 b) + \mu(ab) - \mu(b) \\
&= \mu(a_1) + \mu(ab) - \mu(a_1 b) = \mu(a_1 + ab).
\end{aligned}$$

Thus we have $(b, a)M$. Now consider an element $x \in L$ for which $b \leqq x \leqq a + b$. Clearly $ax \leqq a \leqq c$ and hence $a, ax \in L(\mu)$. It follows that

$$\begin{aligned}
\mu(ax + b) &= \mu(ax) + \mu(b) - \mu(ab) \\
&= \mu(x) + \mu(a) - \mu(a + x) + \mu(b) - \mu(ab) \\
&= \mu(x) + \mu(a) + \mu(b) - \mu(a + b) - \mu(ab) \\
&= \mu(x).
\end{aligned}$$

Since μ is proper we have $x=ax+b$ and this completes the proof.

COROLLARY. *If $p\in L$ is a point, then $p\in L(\mu)$ implies that $b+p$ is prime over b for every $b\in L$ such that $bp=0$.*

PROOF. Consider a point $p\in L(\mu)$ and an element $b\in L$ for which $bp=0$. We always have $(b, p)M$ and thus the first part of our conclusion is vacuous.[7] The second part yields $x=b+px$ for every $x\in L$ for which $b\leqq x\leqq b+p$. Since $px=0$ or $px=p$, it follows immediately that $b+p$ is prime over b.

REMARK. Consider the example of a lattice of six elements 0, p, q, b, c, 1, with $p<c$, $q<b$, 0 the least element, and 1 the greatest element. Our corollary yields the fact that no proper real valued function μ can be defined over this lattice for which p or q is μ-measurable.

LEHIGH UNIVERSITY

[7] L. R. Wilcox, op. cit., p. 491.

A STIELTJES INTEGRAL EQUATION[1]

F. G. DRESSEL

Introduction. By making use of the notion of derivative taken with respect to a function of bounded variation, solutions of the Young-Stieltjes integral equations of the following type

$$F(x) = M(x) + \lambda \int_0^1 H(x, y)dF(y) \tag{1}$$

are obtained in this paper. Since the integration by parts formula is not valid for Young-Stieltjes integrals, equation (1) cannot be immediately changed to the equation

$$f(x) = m(x) + \lambda \int_0^1 f(y)dK(x, y). \tag{2}$$

Fischer[2] in his treatment of equation (2) imposed three conditions on $K(x, y)$ to obtain solutions. In the second part of this paper, we show that one of these conditions is sufficient to insure the existence of a bounded monotone increasing function $g(y)$ such that

$$| K(x, y_2) - K(x, y_1) | \leqq | g(y_2) - g(y_1) |. \tag{3}$$

It has been shown[3] that if condition (3) holds, then equation (2) can be changed into an ordinary Fredholm equation. Thus it appears that only one of the three conditions imposed on $K(x, y)$ by Fischer is needed to insure the solution of (2).

Before handling equation (1), let us note that if $g(y_1) < g(y_2)$ for $y_1 < y_2$, we may apply the transformation[4]

$$\beta(s) = \limsup E_y[s \geqq g(y)], \qquad g(0) \leqq s \leqq g(1), \tag{4}$$

and send the integral equation

$$v(x) = w(x) + \lambda \int_0^1 G(x, y)v(y)dg(y) \tag{5}$$

into the Fredholm equation

[1] Presented to the Society, February 24, 1940.

[2] C. A. Fischer, *The Fredholm theory of the Stieltjes equation*, Annals of Mathematics, (2), vol. 25 (1923–1924), pp. 124–158.

All functions used in this paper are assumed to be measurable Borel.

[3] F. G. Dressel, *A note on Fredholm-Stieltjes integral equations*, this Bulletin, vol. 44 (1938), pp. 434–437.

[4] $E_y[s \geqq g(y)]$ designates the set of values of y such that $s \geqq g(y)$.

$$v(\beta(t)) = w(\beta(t)) + \lambda \int_{\alpha}^{\delta} G(\beta(t), \beta(s))v(\beta(s))ds, \tag{6}$$

where $\alpha = g(0)$, $\delta = g(1)$, and t is any of the possible solutions of

$$x = \beta(t), \qquad g(0) \leqq t \leqq g(1).$$

Now in L^2 the Fredholm theory applies to (6) if

$$\int_0^1 \int_0^1 |G(x, y)|^2 dg(x)dg(y) = \int_{\alpha}^{\delta} \int_{\alpha}^{\delta} |G(\beta(t), \beta(s))|^2 dtds \tag{7}$$

exists.[5] Also if $h(t) \subset L^2$, we have by Schwartz's inequality

$$\left| \int_0^1 G(x, \beta(t))h(t)dt \right|^2 \leqq \int_0^1 |G(x, \beta(t))|^2 dt \int_0^1 |h(t)|^2 dt,$$

and from this we see $\int_0^1 G(x, \beta(t))h(t)dt$ is integrable with respect to $g(x)$ if (7) holds. The above statements lead to the following lemma:

LEMMA 1. *If $g(x_1) < g(x_2)$ for $x_1 < x_2$ with $g(x)$ bounded, and if under the transformation* (4) *the functions $w(x)$, $G(x, y) \subset L^2$, and also if* (7) *holds, then for all values of λ, except perhaps the characteristic values*[5] *of the integral equation* (6), *the equation* (5) *has a solution integrable with respect to $g(x)$.*

Stieltjes integral equations. Consider now the equation (1) where the given functions $M(x)$ and $H(x, y)$ are assumed to be absolutely continuous[6] $g(x)$. Here $g(x)$ is a bounded increasing function, continuous from the right except perhaps at $x = 0$. Let $m(x)$ and $h(x, y)$ be the derived numbers of $M(x)$ and $H(x, y)$ taken with respect to $g(x)$, and form the equation

$$v(x) = m(x) + \lambda \int_0^1 h(x, y)v(y)dg(y). \tag{8}$$

We see from our lemma that equation (8) will have a solution under

[5] For list of references and terminology see E. Hille and J. D. Tamarkin, *On the characteristic values of linear integral equations*, Proceedings of the National Academy of Sciences, vol. 14 (1928), pp. 911–914.

[6] For terminology see A. J. Maria, *Generalized derivatives*, Annals of Mathematics, (2), vol. 28 (1926–1927), pp. 419–432. Let $f(x)$ be a function of bounded variation and denote by $\bar{f}(e)$ the completely additive function of sets defined by $f(x)$. (The facts here are that for any interval $\alpha(x_1 \leqq t \leqq x_2)$ the relation $\bar{f}(\alpha) = f(x_2 + 0) - f(x_1 - 0)$ holds.) The function $f(x)$ is said to be absolutely continuous with respect to the non-decreasing function $g(x)$ if whenever E is a set for which $\bar{f}(E)$ is defined and $\bar{g}(E) = 0$, then $\bar{f}(E) = 0$.

rather weak restrictions on $m(x)$ and $h(x, y)$. Let us now assume (8) has a solution, say $\bar{v}(x)$, which is integrable $g(x)$, and define the function $Q(y)$ as

$$Q(y) = \int_0^y \bar{v}(x)dg(x), \qquad y \neq 0,$$
$$= 0, \qquad y = 0.$$

We are now able to write the solution of (1) as

$$F(x) = M(x) + \lambda \int_0^1 H(x, y)dQ(y). \tag{9}$$

To see that the above $F(x)$ furnishes a solution of (1), substitute the right side of (9) in both sides of equation (1). After cancelling $M(x)$ we have

$$\lambda \int_0^1 H(x, y)dQ(y) = \lambda \int_0^1 H(x, y)dM(y) + \lambda^2 \int_0^1 H(x, y)d_y \int_0^1 H(y, t)dQ(t). \tag{10}$$

Since $\bar{v}(x)$ is a solution of (8), on integration of that equation with respect to $g(x)$, there results

$$Q(x) = M(x + 0) - M(0) + \lambda \int_0^1 [H(x + 0, t) - H(0, t)]\bar{v}(t)dg(t),^7 \quad x \neq 0,$$
$$= M(x + 0) - M(0) + \lambda \int_0^1 [H(x + 0, t) - H(0, t)]dQ(t), \quad x \neq 0,$$
$$= 0, \qquad x = 0.$$

If the above is substituted in the left side of (10), it is easily seen to be equal to the right side, hence (9) satisfies (1).

From the above and our lemma we have the following theorem:

THEOREM 1. *Let $g(x)$ be continuous from the right in $0 < x \leqq 1$, and satisfy the conditions of Lemma* 1; *let $M(x)$ and $H(x, y)$ in equation* (1) *be absolutely continuous $g(x)$ and have $m(x)$ and $h(x, y)$ as derived numbers $(g(x))$. Then if $m(x)$, $h(x, y)$ satisfy the conditions placed on $w(x)$, $G(x, y)$ respectively in Lemma* 1, *equation* (9) *furnishes a solution of* (1) *for all values of λ except those that are characteristic values of the integral equation* (8).

[7] F. G. Dressel, loc. cit., p. 435.

COROLLARY. *If*

$$|H(x_2, y) - H(x_1, y)| \leqq |g(x_2) - g(x_1)|,$$
$$|M(x_2) - M(x_1)| \leqq |g(x_2) - g(x_1)|$$

where $g(x)$ *meets the requirements of Theorem* 1, *then* (9) *is the solution of* (1) *for all values of* λ *that are not characteristic values of the integral equation* (8).

The Fischer theory. In the treatment of equation (2), Fischer imposed the following condition on the kernel $K(x, y) = K(x, y+0)$:

$$V_2K(x, y) < N.$$

Here N is a constant and V_2K has the following meaning:

$$V_2K(x, y) = \text{l.u.b.} \sum_{i=1}^{n} |K(x_i, y_i) - K(x_i, y_{i-1})|,$$

it being understood that $0 = y_0 < y_1 < \cdots < y_n = 1$, and $\{x_i\}$ are any x's in $0 \leqq x \leqq 1$. We shall soon show that the above condition implies that inequality (3) holds.

Define the set E on the interval $0 \leqq y \leqq 1$ to be given by

$$y = m/2^n, \qquad m = 0, 1, 2, \cdots, m \leqq 2^n, \ n = 1, 2, \cdots,$$

and let $g(x, y)$ be a function with the properties

(a) $$g(x, y) = g(y, x) \geqq 0, \qquad g(x, y) \leqq g(x, z) + g(z, y)$$

and if $0 \leqq y_0 < y_1 < \cdots < y_n \leqq 1$ then

(b) $$\sum_{i=1}^{n} g(y_i, y_{i-1}) < N \qquad \text{for all } n.$$

Let $y_j \subset E$ and define $A_n(y_j)$ as

$$A_n(y_j) = \sum_{i=1}^{m} g\left(\frac{i}{2^n}, \frac{i-1}{2^n}\right), \qquad 0 < \frac{m}{2^n} \leqq y_j < \frac{m+1}{2^n},$$
$$= 0, \qquad y_j < \frac{1}{2^n}.$$

Due to the conditions on $g(x, y)$ we have

$$A_n(y_j) \leqq A_{n+1}(y_j), \qquad A_n(y_j) < N;$$

hence $\lim_{n \to \infty} A_n(y_j)$ exists: let us call it $A(y_j)$. Again let $y_1, y_2 \subset E$ with $y_1 < y_2$; then if $y_1 = r/2^k$, $y_2 = m/2^k$,

$$A_k(y_2) - A_k(y_1) = \sum_{i=r+1}^{m} g\left(\frac{i}{2^k}, \frac{i-1}{2^k}\right) \geqq g(y_2, y_1);$$

hence

$$A_k(y_2) - A_k(y_1) \geqq g(y_2, y_1) \geqq 0.$$

Thus $A(y)$ as defined on E is a bounded monotone increasing function; hence we may extend the definition of $A(y)$ on CE so as to define a monotone increasing function of bounded variation[8] on $0 \leqq y \leqq 1$. We make $A(y)$ unique by taking $A(y+0) = A(y)$ if $y \subset CE$.

Turn now to Fischer's condition $V_2K(x, y)$ and define

$$g(y_2, y_1) = \text{l.u.b.}_{0 \leqq x \leqq 1} \, | K(x, y_2) - K(x, y_1) |.$$

This $g(x, y)$ is easily seen to satisfy conditions (a) and (b); hence the following theorem is valid.

THEOREM 2. *If for the function $K(x, y) = K(x, y+0)$ the condition $V_2K(x, y) < N$ holds, then there exists a bounded monotone increasing function $A(y)$ such that*

$$| K(x, y_2) - K(x, y_1) | \leqq | A(y_2) - A(y_1) |.$$

COROLLARY. *The condition $V_2K(x, y) < N$ imposed on $K(x, y) = K(x, y+0)$ by Fischer is sufficient to change equation (2) into an ordinary Fredholm equation.*[9]

The existence of the following limit is essential in Fischer's treatment of equation (2):

$$\lim_{\delta \to 0} \sum_{i=1}^{n} [K(x_i, x_i) - K(x_i, x_{i-1})], \qquad 0 = x_0 < x_1 < \cdots < x_n = 1,$$

where $\delta = \max (x_i - x_{i-1})$. The Fischer theory will therefore not apply to the kernel

$$K_1(x, y) = 1, \quad x = \tfrac{1}{2}, \quad \tfrac{1}{2} \leqq y \leqq 1,$$
$$= 0, \quad \text{elsewhere.}$$

However, the integral equation (2) with the above kernel is easily treated since $V_2K_1(x, y) < 2$.

BROWN UNIVERSITY AND
DUKE UNIVERSITY

[8] G. C. Evans, *Logarithmic Potential*, American Mathematical Society Colloquium Publications, vol. 6, 1927, pp. 5–6.

[9] F. G. Dressel, loc. cit., p. 436.

FOURIER COEFFICIENTS OF BOUNDED FUNCTIONS[1]

BERNARD FRIEDMAN

The results of this paper are divided into two parts. First: inequalities for the Fourier coefficients of any bounded function; second: an approximation theorem for the Fourier development of an arbitrary bounded function.

Inequalities for Fourier coefficients have been discussed in a paper by Professor Szász.[2] However, his work deals mainly with linear inequalities for complex coefficients. The inequalities to be investigated in this paper are not linear. Nevertheless, they are the best possible, for this reason: given any set of numbers which makes the inequality an equality, there exists a bounded function which has these numbers as its Fourier coefficients.

A simple illustration will clarify this. Let $f(x)$ be a bounded measurable function in $(-\pi, \pi)$ such that $|f(x)| \leqq 1$. The Fourier coefficients of $f(x)$ are given by the formulae

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx,$$

$$n = 1, 2, 3, \cdots.$$

Then, it is clear that

$$|a_n| \leqq \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)||\cos nx|\,dx \leqq \frac{1}{\pi}\int_{-\pi}^{\pi} |\cos nx|\,dx,$$

$$|b_n| \leqq \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)||\sin nx|\,dx \leqq \frac{1}{\pi}\int_{-\pi}^{\pi} |\sin nx|\,dx$$

since $|f(x)| \leqq 1$.

Since the cosine is negative in the intervals $(-\pi, -\pi/2)$ and $(\pi/2, \pi)$ and positive in the remaining interval $(-\pi/2, \pi/2)$,

$$\int_{-\pi}^{\pi} |\cos x|\,dx = \int_{-\pi}^{-\pi/2} (-\cos x)dx + \int_{-\pi/2}^{\pi/2} \cos x\,dx + \int_{\pi/2}^{\pi} (-\cos x)dx = 4.$$

Therefore, $|a_1| \leqq 4/\pi$. However, if $f_1(x) = -1$, $-\pi < x < -\pi/2$; $f_1(x)$

[1] Presented to the Society, September 8, 1939.

[2] American Journal of Mathematics, vol. 61 (1939).

$=1$, $-\pi/2<x<\pi/2$; $f_1(x)=-1$, $\pi/2<x<\pi$; then

$$\int_{-\pi}^{\pi} f_1(x)\cos x\,dx = \int_{-\pi}^{-\pi/2}(-\cos x)dx + \int_{-\pi/2}^{\pi/2}\cos x\,dx + \int_{\pi/2}^{\pi}(-\cos x)dx$$

and the bound $4/\pi$ will actually be attained for the function $f_1(x)$.

Likewise, $\int_{-\pi}^{\pi}|\cos nx|\,dx$ and $\int_{-\pi}^{\pi}|\sin nx|\,dx$ both equal four, so that $|a_n|$ and $|b_n| \leqq 4/\pi$; these bounds will be attained for the functions signum cos nx and signum sin nx respectively.

But these bounds for the Fourier coefficients cannot be attained simultaneously. If one of the coefficients is too large, the others must be small. For example: b_1 and b_2 must satisfy the following inequality:

$$\left(\frac{\pi}{4}b_1\right)^2 + \frac{\pi}{4}|b_2| \leqq 1$$

so that if $b_1=4/\pi$ then $b_2=0$ while if $b_2=4/\pi$ then $b_1=0$. Conversely, any numbers b_1, b_2 satisfying this relation will be the Fourier coefficients of some bounded function.

Similar relations hold between the other Fourier coefficients, but they are much more complicated:

$$a_n^2 + b_n^2 \leqq \left(\frac{4}{\pi}\right)^2,$$

$$\left(\frac{\pi}{4}b_1+1\right)^3 - 3\left(\frac{\pi}{4}b_1+1\right) - 1 \leqq 3\left(\frac{\pi}{4}b_3\right) \leqq \left(\frac{\pi}{4}b_1-1\right)^3 - 3\left(\frac{\pi}{4}b_1-1\right)+1.$$

After the inequalities for the Fourier coefficients are obtained, they are used to derive an approximation theorem. But instead of being an approximation to the function itself, it will be an approximation to the Fourier development of the function. More precisely: take a bounded measurable function $f(x)$ whose absolute value is less than one; let its Fourier development be $a_0/2+\sum(a_n\cos nx+b_n\sin nx)$. Then, for any N there exists a function $g_N(x)$ which takes the values 1 and -1 only, such that the first N Fourier coefficients of $g_N(x)$ are the same as the first N Fourier coefficients of $f(x)$.

In general, there will be $N+1$ discontinuities. If $f(x)$ is even or odd, $g_N(x)$ will be even or odd. For example, let $f(x)=(\pi-x)/\pi$,

$0<x<\pi; f(-x)=-f(x)$. Then its Fourier series will be

$$\frac{2}{\pi}\sum_{1}^{\infty}\frac{\sin nx}{n},$$

$$g_1(x)=+1, \quad 0<x<\frac{2\pi}{3}; \qquad g_2(x)=+1, \quad 0<x<\frac{\pi}{2};$$

$$g_1(x)=-1, \quad \frac{2\pi}{3}<x<\pi; \qquad g_2(x)=-1, \quad \frac{\pi}{2}<x<\frac{2\pi}{3};$$

$$g_2(x)=+1, \quad \frac{2\pi}{3}<x<\pi.$$

Incidentally, one of the inequalities leads to the following result: If $F(x)$ is a bounded measurable function in absolute value less than $\frac{1}{2}$,

$$\int_0^{\pi} F(x)\cos x\,dx \leqq \cos\left(\int_0^{\pi} F(x)dx\right).$$

The method used in proving the above inequalities is as follows: Let $f(x)$ be a bounded measurable function in $(-\pi, \pi)$ such that $|f(x)| \leqq 1$. Let the Fourier development of $f(x)$ be

$$\frac{a_0}{2}+\sum (a_n\cos nx + b_n \sin nx).$$

Consider the expression

$$(1) \qquad f_n(\mu, a) \equiv \mu_0 a_0 + \mu_1 b_1 + \mu_2 a_1 + \cdots + \mu_{2n-1} b_n + \mu_{2n} a_n$$

where $\mu_0, \mu_1, \cdots, \mu_{2n}$ are arbitrary real parameters. The functions $f_n(\mu, a)$ can be expressed in terms of the function $f(x)$ by using the usual formulae for Fourier coefficients

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx, \qquad b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx.$$

From these formulae and (1), it is clear that

$$f_n(\mu, a)=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\left[\mu_0+\mu_1\sin x+\mu_2\cos x+\cdots + \mu_{2n-1}\sin nx+\mu_{2n}\cos nx\right]dx$$

$$\equiv \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)C_n(\mu, x)dx,$$

say.

Assuming the μ's are fixed, the maximum of $f_n(\mu, a)$ will be attained when the integrand $f(x)C_n(\mu, x)$ is as large as possible. But

$$|f(x)C_n(\mu, x)| \leqq |C_n(\mu, x)|$$

since $|f(x)| \leqq 1$. However, if $f(x)$ is taken as signum $C_n(\mu, x)$; that is, $f(x)=1$ when $C_n(\mu, x)>0$ and $f(x)=-1$ when $C_n(\mu, x)<0$, then

$$|f(x)C_n(\mu, x)| = |C_n(\mu, x)|.$$

The following result is now obvious:

$$f_n(\mu, a) \leqq \frac{1}{\pi}\int_{-\pi}^{\pi} |C_n(\mu, x)|\, dx \equiv D_n(\mu), \tag{2}$$

and the equality holds only when

$$f(x) = \text{signum } C_n(\mu, x) \equiv g_n(\mu, x).$$

For example, if $n=1$ and $\mu_1=0$, formula (2) states that

$$\mu_0 a_0 + \mu_2 a_1 \leqq \frac{1}{\pi}\int_{-\pi}^{\pi} f_1(x)C_1(\mu, x)dx$$

where

$$f_1(x) = \text{signum } (\mu_0 + \mu_2 \cos x).$$

Assume $\mu_0+\mu_2 \cos x=0$ has the real roots $\pm\alpha_1$; then $g_1(\mu, x)$ will be defined as follows:

$$g_1(\mu, x) = +1, \quad -\alpha_1 < x < \alpha_1,$$

$$g_1(\mu, x) = -1, \quad \begin{cases} -\pi < x < -\alpha_1, \\ \alpha_1 < x < \pi, \end{cases} \quad \text{if } \mu_0 > 0;$$

$$g_1(\mu, x) = -1, \quad -\alpha_1 < x < \alpha_1,$$

$$g_1(\mu, x) = +1, \quad \begin{cases} -\pi < x < -\alpha_1, \\ \alpha_1 < x < \pi, \end{cases} \quad \text{if } \mu_0 < 0;$$

and

$$\int_{-\pi}^{\pi} |C_1(\mu, x)|\, dx = \pm\frac{1}{\pi}\int_{-\pi}^{\pi} g_1(\mu, x)C_1(\mu, x)dx$$

$$= \pm\frac{1}{\pi}\left[\int_{-\pi}^{-\alpha_1} -C_1(\mu, x)dx + \int_{-\alpha_1}^{\alpha_1} C_1(\mu, x)dx + \int_{\alpha_1}^{\pi} -C_1(\mu, x)dx\right]$$

$$= \pm \frac{1}{\pi} [\mu_0(4\alpha - 2\pi) + \mu_2 4 \sin \alpha] = \pm \frac{4}{\pi}\left[\mu_0\left(\alpha - \frac{\pi}{2}\right) + \mu_2 \sin \alpha\right].$$

If μ_0 is positive, the upper sign holds; if μ_0 is negative, the lower sign is used.

But $C_1(\mu, \alpha_1) = 0$; that is, $\mu_0 + \mu_2 \cos \alpha_1 = 0$ or

$$\cos \alpha = -\mu_0/\mu_2,$$

$$\alpha = \arccos(-\mu_0/\mu_2), \qquad \alpha - \pi/2 = \arcsin \mu_0/\mu_2,$$

$$\sin \alpha = (1 - \mu_0^2/\mu_2^2)^{1/2},$$

so that

$$D_1(\mu) = \pm \frac{4}{\pi} [\mu_0 \arcsin(\mu_0/\mu_2) + \mu_2(1 - \mu_0^2/\mu_2^2)^{1/2}].$$

Therefore

$$\mu_0 a_0 + \mu_2 a_1 \leqq \pm \frac{4}{\pi} [\mu_0 \arcsin(\mu_0/\mu_2) + \mu_2(1 - \mu_0^2/\mu_2^2)^{1/2}] \tag{3}$$

where a_0 and a_1 are the Fourier coefficients of any bounded function. It is easy to see that the equality sign holds only when a_0 and a_1 equal the Fourier coefficients of $g_1(\mu, x)$.

Similar results hold for any n. Asume $C_n(\mu, x) = 0$ has $2n$ real roots α_i where $-\pi < \alpha_1 \leqq \alpha_2 \leqq \alpha_3 \leqq \cdots \leqq \alpha_{2n} \leqq \pi$. Call $-\pi = \alpha_0$ and $\pi = \alpha_{2n+1}$.

Assume $C_n(\mu, x)$ is positive between α_0 and α_1. Then $g_n(\mu, x)$ will be defined as follows: $g_n(\mu, x) = (-1)^i$, $\alpha_i < x < \alpha_{i+1}$, $i = 0, 1, \cdots, 2n+1$. If $C_n(\mu, x)$ is negative, then $g_n(\mu, x) = (-1)^{i+1}$, $\alpha_i < x < \alpha_{i+1}$. Therefore,

$$D_n(\mu) = \frac{1}{\pi} \int_{-\pi}^{\pi} g_n(\mu, x) C_n(\mu, x) dx = \pm \frac{1}{\pi} \sum_0^{2n+1} \int_{\alpha_1}^{\alpha_{i+1}} (-1)^i C_n(\mu, x) dx$$

so that from (2)

$$\mu_0 a_0 + \mu_1 b_1 + \mu_2 a_1 + \cdots + \mu_{2n-1} b_n + \mu_{2n} a_n \leqq D_n(\mu). \tag{4}$$

Let

$$a_i = x_{2i}, \qquad b_i = x_{2i-1}. \tag{5}$$

Consider the $2n+1$ dimensional space of points whose coordinates are the x_j ($j = 0, 1, 2, \cdots, 2n+1$). The inequality (4) can be written as

$$\sum_{j=0}^{2n+1} \mu_j x_j \leqq D_n(\mu).$$

This implies that, for any bounded measurable function, the point P whose coordinates are given by (5) must lie on one side of the hyperplane whose equation is

$$\sum_{0}^{2n+1} \mu_j x_j = D_n(\mu). \tag{6}$$

The point P will lie on the plane if (4) becomes an equality. This will occur when and only when $f(x)$ has its first $2n+1$ Fourier coefficients equal to those of $g_n(\mu, x)$.

Since the μ_j were arbitrary, the same result holds for any set $\mu_0, \mu_1, \cdots, \mu_{2n+1}$. Therefore, the point P, whose coordinates are (x_j), $j=0, 1, \cdots, 2n+1$, must lie on one side of every plane in the family of planes defined by the equation (6).

The previous example will clarify this. Equation (3) implies that if, for any $f(x)$, a_0 is plotted as the x coordinate and a_1 as the y coordinate of a point, then such points will lie on the origin side of the straight line whose equation is

$$\mu_0 x + \mu_2 y = D_1(\mu).$$

The point will lie on the straight line if and only if $f(x)$ has its Fourier coefficients a_0 and a_1 equal to those of $g_1(\mu, x)$.

Since the μ's were arbitrary, the same result holds for any set μ_0, μ_2. Therefore, the point P, whose coordinates are (a_0, a_1) must lie on one side of every line in the family of straight lines defined by the equation

$$\mu_0 x + \mu_2 y = D_1(\mu).$$

This implies that the point P will lie inside the region enveloped by these straight lines.

The usual procedure in finding the equation of this envelope is, first, to differentiate partially with respect to the parameter, and second, to eliminate the parameters between the resulting equation and the original equation. The partial differentiation results in the following two equations:

$$x = \frac{\partial D}{\partial \mu_0} = \pm (4/\pi) \arcsin (\mu_0/\mu_2), \quad y = \frac{\partial D}{\partial \mu_2} = \pm (4/\pi)(1 - \mu_0^2/\mu_2^2)^{1/2}.$$

After the elimination of the parameters, the following equation is obtained for the envelope:

$$y = (4/\pi) \cos (\pi x/4).$$

This implies that

$$a_1 \leqq (4/\pi) \cos (\pi a_0/4).$$

Equality will hold, as before, if $f(x)$ has the same first two Fourier coefficients as any function $g_1(\mu, x)$.

In the general case, the equation of the envelope is obtained by the same method. When formula (6) is differentiated partially with respect to the parameters, the following equations result:

$$x_j = \frac{\partial D_n(\mu)}{\partial \mu_j}, \qquad j = 0, 1, \cdots, 2n+1.$$

But

$$\frac{\partial D_n(\mu)}{\partial \mu_j} = \pm \frac{1}{\pi} \frac{\partial}{\partial \mu_j}\left[\sum_{i=0}^{2n+1} \int_{\alpha_i}^{\alpha_{i+1}} (-1)^i C_n(\mu, x)dx \right]$$

where the α_i, the roots of $C_n(\mu, x) = 0$, also depend on the μ_j. Now

$$\frac{\partial}{\partial \mu_j} \int_{\alpha_i}^{\alpha_{i+1}} (-1)^i C_n(\mu, x)dx = \int_{\alpha_i}^{\alpha_{i+1}} (-1)^i \frac{\partial C_n(\mu, x)}{\partial \mu_j} dx$$
$$+ (-1)^i C_n(\mu, \alpha_{i+1}) \frac{\partial \alpha_{i+1}}{\partial \mu_j} - (-1)^i C_n(\mu, \alpha_i) \frac{\partial \alpha_i}{\partial u_j}$$

by the standard rule for differentiating under the integral sign.

However $C_n(\mu, \alpha_{i+1}) = C_n(\mu, \alpha_i) = 0$ and

$$\frac{\partial C_n(\mu, x)}{\partial \mu_j} = \sin (j+1)x/2 \text{ if } j \text{ is odd},$$
$$= \cos (jx/2) \text{ if } j \text{ is even},$$

so that

$$x_j = \pm \frac{1}{\pi} \sum_{i=0}^{2n+1} \int_{\alpha_i}^{\alpha_{i+1}} (-1)^i \sin (j+1)x/2dx \text{ if } j \text{ is odd},$$
$$= \pm \frac{1}{\pi} \sum_{i=0}^{2n+1} \int_{\alpha_i}^{\alpha_{i+1}} (-1)^i \cos (jx/2)dx \text{ if } j \text{ is even}, \tag{7}$$

where the signs are either all positive or all negative.

Instead of eliminating the parameters μ, it is simpler to eliminate the parameters α, thus obtaining the equation of the envelope $\phi_n(x_j) = 0$ which implies the inequality

$$\phi_n(a, b) \leqq 0. \tag{8}$$

Since $\phi_n(x_j)$ is obtained as the envelope of planes, it is obvious that

the region bounded by it is the convex hull of a family of half-spaces. Again, the inequality will hold if and only if $f(x)$ has its first $2n+1$ Fourier coefficients $(a_0, a_1, b_1, \cdots, a_n, b_n)$ equal to the corresponding Fourier coefficients of any one of the functions $g_n(\mu, x)$.

A few illustrations will clarify this method. Let $n=2$ and $\mu_0=\mu_2=\mu_4=0$. Then the α_i will be the roots of the equation

$$\mu_1 \sin x + \mu_3 \sin 2x = 0,$$

and from formula (7)

$$x_1 = \pm \frac{1}{\pi} \sum \int_{\alpha_i}^{\alpha_{i+1}} (-1)^i \sin x \, dx = \pm (-4/\pi \cos \alpha_3),$$

since $\alpha_1 = -\alpha_3$, and

$$x_1 = \pm \frac{1}{\pi} \sum \int_{\alpha_i}^{\alpha_{i+1}} (-1)^i \sin 2x \, dx = \pm (2/\pi)(1 - \cos 2\alpha_3).$$

Therefore $(\pi x_1/4)^2 + \pi x_3/4 = 1$ or $(\pi x_1/4)^2 - \pi x_3/4 = 1$ according as the plus or minus signs are used. But since x_1 was the coordinate for b_1 and x_3 for b_2, the following inequality results:

$$(\pi b_1/4)^2 + \pi \left| b_3 \right| /4 \leqq 1.$$

From the result proved above: that the Fourier coefficients of $f(x)$ satisfy certain inequalities or, stating it geometrically, lie in certain convex regions, the approximation theorem follows. The theorem can be stated as follows:

Let $f(x)$ be a real, bounded, measurable function in absolute value less than one. Let its Fourier development be

$$\frac{a_0}{2} + \sum (a_n \cos nx + b_n \sin nx).$$

Then for every N, there exists a step-function $g_N(x)$ which takes on the values of $+1$ and -1 only, such that the first N Fourier coefficients of $g_N(x)$ are the same as the first N Fourier coefficients of $f(x)$, that is,

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g_N(x) \cos nx \, dx, \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g_N(x) \sin nx \, dx,$$

$$n = 0, 1, \cdots, N.$$

Again, let $a_i = x'_{2i}$, $b_i = x'_{2i-1}$. Then, as before, the point P' whose coordinates are $(x'_0, x'_1, \cdots, x'_{2N})$ must satisfy the inequality (8):

$\phi_N(x_j') \leqq 0$, or, geometrically, the point P' must lie inside the convex region bounded by the hypersurface whose equation is

$$\phi_N(x_0, x_1, x_2, \cdots, x_{2N}') = 0.$$

Suppose now that P' lies on the hypersurface, that is, its coordinates satisfy the equation

$$\phi_N(x_0', x_1', x_2', \cdots, x_{2N}') = 0.$$

Since $\phi_N(x_j) = 0$ was obtained as the envelope of the family of hyperplanes represented by equation (6), $\sum \mu_j x_j = D_N(\mu)$, there will be a plane of the family tangent to the surface at P'. But since P' lies on this plane, whose equation can be written

$$\mu_0' x_0 + \mu_1' x_1 + \mu_2' x_2 + \cdots + \mu_{2N}' x_{2N} = D_N(\mu'),$$

its coordinates must equal those of the function $g_N(\mu', x)$. Therefore, $g_N(\mu_0', \cdots, \mu_{2N}', x)$ is the function described in the theorem.

If P' lies inside the hypersurface, consider the points in $2N+3$ dimensional space whose coordinates are $(x_0', x_1', x_2', \cdots, x_{2N}', x_{2N+1}, x_{2N+2})$ where x_{2N+1} is an arbitrary variable. They will lie on the straight line through $(x_0', x_1', x_2', \cdots, x_{2N}', 0, 0)$ perpendicular to the hyperplane $x_{2N+1} = x_{2N+2} = 0$.

Since $\phi_{N+1}(x_j) = 0$ is a closed convex surface for all n, the above straight line will intersect it in some point whose coordinates are $(x_0', x_1', x_2', \cdots, x_{2N}', x_{2N+1}', x_{2N+2}')$. Just as before, there will exist a step-function $g_{N+1}(x)$ whose Fourier development is $a_0'/2 + \sum (a_n' \cos nx + b_n' \sin nx)$ and such that $a_i' = x_{2i}'$, $b_i' = x_{2i-1}'$; $i = 0, 1, 2, \cdots, N+1$.

This function satisfies the conditions of the theorem for

$$x_{2i}' = a_i = a_i', \qquad x_{2i-1}' = b_i' = b_i.$$

UNIVERSITY OF WISCONSIN

THE GENERALIZATION OF A LEMMA OF M. S. KAKEYA

J. GERONIMUS

We shall prove the following:

LEMMA. *It is always possible to find the unique polynomial*

$$\phi^*(z) = \sum_{k=0}^{2s} \gamma_k^* z^k$$

of degree $2s$ *possessing the following properties*:

I. $$\phi^*(z) = ci^2(z)\tau(z)\tau^*(z), \qquad c = \text{const.},$$

the polynomial $i(z)$ *of degree* $\sigma \leqq s$ *having all roots in the domain* $|z| > 1$:

$$i(z) = \prod_{i=1}^{\sigma} (z - a_i), \qquad |a_i| > 1, \quad i = 1, 2, \cdots, \sigma,$$

and the polynomial $\tau(z)$ *being of degree* $\nu = s - \sigma$:

$$\tau(z) = \prod_{i=1}^{\nu} (z - \alpha_i), \qquad \tau^*(z) = z^\nu \bar{\tau}\left(\frac{1}{z}\right) = \prod_{i=1}^{\nu} (1 - z\bar{\alpha}_i).$$

II. *It is subject to the conditions*

$$\omega_i(\phi^*) = \sum_{k=0}^{2s} \gamma_k^* c_k^{(i)} = d_i, \qquad i = 0, 1, \cdots, s,$$

the given linear functionals ω_i *being such that every polynomial* $\phi(z)$ *of degree* $n \geqq 2s$ *for which*

$$\omega_i(\phi) = \sum_{k=0}^{2s} \gamma_k c_k^{(i)} = 0, \quad (i = 0, 1, \cdots, s), \qquad \phi(z) = \sum_{k=0}^{n} \gamma_k z^k,$$

has $s+1$ *roots at least in the domain* $|z| < 1$.

In the particular case when

$$\omega_i(\phi) = \phi^{(i)}(z_k), \qquad |z_k| < 1,$$

this lemma has been proved by M. S. Kakeya [1];[1] without being aware of his result we have proved this lemma in the case[2]

[1] Numbers in brackets refer to the bibliography at the end.

[2] In [1] and [2] one may find the application of this lemma to some extremal problems.

$$\omega_i(\phi) = \frac{1}{i!}\left(\frac{d^i\phi}{dz^i}\right)_{z=0}, \qquad i = 0, 1, \cdots, s.$$

In order to prove this lemma in the most general case we consider the following extremal problem:

PROBLEM. *To find the minimum of the integral*

$$L(b) = \int_0^{2\pi} |t(z)|^2 b(\theta)\,d\theta, \qquad z = e^{i\theta}, \tag{1}$$

$t(z)$ being the given polynomial of degree s with $t(0)\neq 0$ and $b(\theta)$ being a trigonometric polynomial of order $n\geqq 2s$:

$$b(\theta) = R\left\{z^n\bar{\phi}\left(\frac{1}{z}\right)\right\} = R\sum_{k=0}^{n}\bar{\gamma}_k e^{i(n-k)\theta}, \qquad z = e^{i\theta},$$

subject to the conditions[3]

$$\omega_i(b) = \omega_i(\phi) = \sum_{k=0}^{2s}\gamma_k c_k^{(i)} = d_i, \qquad i = 0, 1, \cdots, s. \tag{2}$$

The fundamental property of our functionals ω_i yields at once that every trigonometric polynomial $b(\theta)$ subject to the conditions

$$\omega_i(b) = 0, \qquad i = 0, 1, \cdots, s,$$

has in $(0, 2\pi)$ no more than $2(n-s-1)$ changes of sign. It is clear that there exists a solution of our problem. Further, the necessary conditions for an extremum are

$$\operatorname{sgn} b^*(\theta)\,|t(z)|^2 = R\sum_{k=n-2s}^{\infty}A_k z^k, \qquad z = e^{i\theta},$$

whence we find at once that the Fourier expansion of $\operatorname{sgn} b^*(\theta)$ is of the form

$$\operatorname{sgn} b^*(\theta) = R\sum_{k=n-s}^{\infty}B_k z^k, \qquad z = e^{i\theta}.$$

We have shown in [2] that every trigonometric polynomial with this property must be of the form

$$b^*(\theta) = R\{\bar{c}z^{n-2s+\nu}q^2(z)\}\tau(z)\bar{\tau}(1/z), \qquad z = e^{i\theta},$$

$q(z)$ being a polynomial of degree $\sigma\leqq s$ all of whose roots lie in the domain $|z|<1$, and $\tau(z)$ being a polynomial of degree $\nu=s-\sigma$.

[3] The functionals ω_i are the same as above.

The polynomial $b^*(\theta)$ for which the minimum is attained is unique. If there were two such polynomials, $b_1^*(\theta)$ and $b_2^*(\theta)$, then we would have

$$L(b_1^*) \leqq L\left(\frac{b_1^* + b_2^*}{2}\right) \leqq L(b_1^*);$$

then $b_1^*(\theta)$ and $b_2^*(\theta)$ would change sign at the same points, that is, the polynomial

$$b_1^*(\theta) - b_2^*(\theta) = R\{z^{n-2s+\nu}q^2(z)\}\{\bar{c}_1|\tau_1(z)|^2 - \bar{c}_2|\tau_2(z)|^2\}, \quad z = e^{i\theta},$$

would have at least $2(n-\nu)$ changes of sign in $(0, 2\pi)$; but since

$$\omega_i(b_1^* - b_2^*) = 0, \qquad i = 0, 1, \cdots, s,$$

the polynomial $b_1^*(\theta) - b_2^*(\theta)$ cannot have more than $2(n-s-1)$ changes of sign in $(0, 2\pi)$; this contradiction proves the unicity of the polynomial solving our problem. *Thus we find that there exists the unique polynomial* $b^*(\theta)$ *minimizing* (1) *under conditions* (2) *and it must be of the form*

$$\begin{aligned} b^*(\theta) &= R\{\bar{c}z^{n-2s+\nu}q^2(z)\tau(z)\bar{\tau}(1/z)\} \\ &= R\{\bar{\gamma}_0^* z^n + \bar{\gamma}_1^* z^{n-1} + \cdots + \bar{\gamma}_{2s}^* z^{n-2s}\}, \qquad z = e^{i\theta}. \end{aligned}$$

Since the real parts of two polynomials coincide on the unit circle, these polynomials are identical, that is,

$$\bar{c}z^{n-2s}q^2(z)\tau(z)\tau^*(z) = \bar{\gamma}_0^* z^n + \bar{\gamma}_1^* z^{n-1} + \cdots + \bar{\gamma}_{2s}^* z^{n-2s},$$

whence we find finally

$$\phi^*(z) = \gamma_0^* + \gamma_1^* z + \cdots + \gamma_{2s}^* z^{2s} = ci^2(z)\tau(z)\tau^*(z),$$

where

$$i(z) = q^*(z) = z^\nu \bar{q}(1/z).$$

Thus we have found the polynomial $\phi^*(z)$ satisfying all the conditions of our lemma.

Bibliography

1. S. Kakeya, *Maximum modulus of some expressions of limited analytic functions*, Transactions of this Society, vol. 22 (1921), pp. 489–504.

2. J. Geronimus, *On a problem of F. Riesz and on the generalized problem of Tchebycheff-Korkine-Zolotareff*, Bulletin de l'Académie des Sciences de l'URSS., vol. 3 (1939), pp. 279–288.

Technological Institute,
Kharkow, U.S.S.R.

NOTE ON CERTAIN ORTHOGONAL POLYNOMIALS[1]

DUNHAM JACKSON

1. **Introduction.** A well known theorem[2] states that if $K_n(x, t) = \sum_0^n p_k(x)p_k(t)$ is the kernel associated with a system of orthonormal polynomials on an interval (a, b) with weight function $\rho(t)$, and if x_0 is a real number not belonging to the open interval (a, b), the function $K_n(x, t)$ is orthogonal to every polynomial of lower degree with respect to $|t-x_0|\rho(t)$ as weight function. This result is to be extended below to orthogonal trigonometric sums, and more generally to other orthogonal polynomials in two real variables on an algebraic curve.[3] The fact that a single polynomial of the nth degree in the original formulation is replaced by two or more sums or polynomials of like degree in the generalized orthogonal systems imparts to the extension some features of novelty. Certain other generalizations are briefly mentioned also.

2. **Trigonometric sums.** Let $U_0(x)$, $U_1(x)$, $V_1(x)$, $\cdots$ be orthonormal trigonometric sums for weight $\rho(x)$, and let

$$K_n(x, s) = U_0(x)U_0(s) + U_1(x)U_1(s) + \cdots + U_n(x)U_n(s) + V_1(x)V_1(s) + \cdots + V_n(x)V_n(s).$$

For definiteness it may be assumed that $U_k(x)$ contains no term in $\sin kx$, while $V_k(x)$ contains $\sin kx$ with a nonvanishing coefficient; the function $K_n(x, s)$ would be unchanged if U_k, V_k were replaced by any equivalent pair of sums of the kth order. If $T_n(x)$ is any trigonometric sum of the nth order,

$$\int_{-\pi}^{\pi} \rho(s)K_n(x, s)T_n(s)ds = T_n(x).$$

Suppose $\rho(x) \equiv 0$ throughout an interval (α, β). Let x_1, x_2 be any two distinct points of (α, β). Let $\tau_{n-1}(x)$ be an arbitrary trigonometric sum of order $n-1$ at most, and let $\sin \frac{1}{2}(x-x_1) \sin \frac{1}{2}(x-x_2)\tau_{n-1}(x) = T_n(x)$; the product

$$\sin \tfrac{1}{2}(x - x_1) \sin \tfrac{1}{2}(x - x_2) = \tfrac{1}{2} \cos \tfrac{1}{2}(x_1 - x_2) - \tfrac{1}{2} \cos [x - \tfrac{1}{2}(x_1 + x_2)]$$

---

[1] Presented to the Society, December 29, 1939.

[2] See, e.g., G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. 23, New York, 1939, p. 39.

[3] See D. Jackson, *Orthogonal polynomials on plane curves*, Duke Mathematical Journal, vol. 3 (1937), pp. 228–236.

is a trigonometric sum of the first order. Then

$$(1)\qquad \int_{-\pi}^{\pi} \rho(s)K_n(x, s) \sin \tfrac{1}{2}(s - x_1) \sin \tfrac{1}{2}(s - x_2)\tau_{n-1}(s)ds = \sin \tfrac{1}{2}(x - x_1) \sin \tfrac{1}{2}(x - x_2)\tau_{n-1}(x).$$

The right-hand member vanishes for $x=x_1$ and for $x=x_2$; that is to say, $K_n(x_1, s)$ *and* $K_n(x_2, s)$ *as trigonometric sums of the nth order in* s *are orthogonal to every sum* $\tau_{n-1}(s)$ *of lower order, with respect to* $\sin \frac{1}{2}(s-x_1) \sin \frac{1}{2}(s-x_2)\rho(s)$ *as weight function.* This weight function is of constant sign, since $\rho(s)$ vanishes throughout the interval (x_1, x_2).

The sums $K_n(x_1, s)$ *and* $K_n(x_2, s)$ *are linearly independent.* An identity $K_n(x_2, s) = CK_n(x_1, s)$ would require in particular that $U_0(x_2) = CU_0(x_1)$, $U_1(x_2) = CU_1(x_1)$, $V_1(x_2) = CV_1(x_1)$, that is, that $C=1$, $\cos x_2 = \cos x_1$, $\sin x_2 = \sin x_1$.

In the limiting case $x_2 = x_1$,

$$\int_{-\pi}^{\pi} \rho(s)K_n(x, s) \sin^2 \tfrac{1}{2}(s - x_1)\tau_{n-1}(s)ds = \sin^2 \tfrac{1}{2}(x - x_1)\tau_{n-1}(x)$$

and by differentiation with respect to x under the sign of integration

$$\int_{-\pi}^{\pi} \rho(s)K_{nx}(x, s) \sin^2 \tfrac{1}{2}(s - x_1)\tau_{n-1}(s)ds = \sin \tfrac{1}{2}(x - x_1) \cos \tfrac{1}{2}(x - x_1)\tau_{n-1}(x) + \sin^2 \tfrac{1}{2}(x - x_1)\tau'_{n-1}(x),$$

where $K_{nx}(x, s)$ means $(\partial/\partial x)K_n(x, s)$. Both right-hand members vanish for $x=x_1$; $K_n(x_1, s)$ and $K_{nx}(x_1, s)$ as trigonometric sums of the nth order in s are orthogonal to every sum of lower order for weight $\sin^2 \frac{1}{2}(s-x_1)\rho(s)$. Here no assumption with regard to vanishing of ρ is required. It is obvious that K_n and K_{nx} are linearly independent, for one contains a constant term and the other does not.

Vanishing of ρ is again unnecessary if x_1 and x_2 are allowed to be complex. Let $x_1 = \gamma + i\delta$, $x_2 = \gamma - i\delta$, with γ and δ real,

$$\sin \tfrac{1}{2}(x - x_1) \sin \tfrac{1}{2}(x - x_2) = \tfrac{1}{2} \cos i\delta - \tfrac{1}{2} \cos (x - \gamma) = \tfrac{1}{2} \cosh \delta - \tfrac{1}{2} \cos (x - \gamma).$$

Since each member of (1) represents a trigonometric sum in x, and they are identical for real values of x, they are equal also when x is complex. For $x=x_1$ the right-hand member is zero, and the real and pure imaginary parts of the expression on the left must vanish separately. The same result is obtained by setting $x=x_2$. The real and pure imaginary parts of $K_n(x_1, s)$ for real s (or of its conjugate $K_n(x_2, s)$) are

orthogonal to every trigonometric sum of lower order in s with respect to the non-negative weight function $[\frac{1}{2}\cosh\delta-\frac{1}{2}\cos(s-\gamma)]\rho(s)$. These real and pure imaginary components, being linearly independent combinations of the linearly independent functions $K_n(x_1, s)$ and $K_n(x_2, s)$, are themselves linearly independent trigonometric sums of the nth order in s.

3. **Polynomials on an algebraic curve.** Orthogonal trigonometric sums can be regarded as a special instance of orthogonal polynomials in two real variables on an algebraic curve, the curve in this case being a circle (or, without essentially greater generality, an ellipse). Considerations analogous to those outlined above apply to the more general problem. As an illustration the reasoning will be formulated for a curve of the third degree. It will be immediately apparent that the same type of argument can be applied to curves of higher degree, if exceptional cases are avoided, or, on the other hand, to an alternative treatment of orthogonal polynomials on a circle, or to orthogonal polynomials on hyperbolic arcs. (Orthogonal polynomials on a parabola are essentially orthogonal polynomials in a single variable.)

Let C be a curve of the third degree. Let $\rho(x, y)$ be a non-negative weight function defined on C, and, for simplicity, vanishing identically except on an arc or arcs of finite extent. Let $p_{01}(x, y)$, $p_{11}(x, y)$, $p_{12}(x, y)$, and, for $n\geqq 2$, $p_{n1}(x, y)$, $p_{n2}(x, y)$, $p_{n3}(x, y)$ denote the members of a corresponding system of orthonormal polynomials on C, each polynomial being of total degree indicated by its first subscript. Let

$$K_n(x, y, u, v) = \sum_{k\leqq n} p_{kj}(x, y)p_{kj}(u, v),$$

the summation being extended over all the relevant pairs of subscripts. As in the special case of trigonometric sums, the function K_n is completely determinate when the points (x, y) and (u, v) are on the curve C in their respective planes, although the individual polynomials of any specified degree in the system admit orthogonal transformation among themselves.

Let $ax+by+c=0$ be the equation of a line intersecting C in three distinct points (x_1, y_1), (x_2, y_2), (x_3, y_3), and so situated that the domain of orthogonality of the polynomials is all on one side of it, that is, so that $ax+by+c$ is of constant sign where $\rho\neq 0$. The intersections may be real, or two of them may be conjugate complex; it is understood that the coefficients a, b, c, as well as those in the equation of C, are real. It is to be understood further, if C is degenerate, that the domain of orthogonality includes a portion of positive measure

(with respect to arc length as variable of integration) on each of the loci defined by the vanishing of an irreducible factor in the equation of C.

If $P_n(x, y)$ is any polynomial of the nth or lower degree, then for (x, y) on C

$$\int_C \rho(u, v)K_n(x, y, u, v)P_n(u, v)ds = P_n(x, y),$$

where s denotes arc length on the curve C in the (u, v) plane. While the identity holds in general only on the curve, it is valid at all points of the curve, not merely where $\rho \neq 0$. If $\pi_{n-1}(x, y)$ is an arbitrary polynomial of degree $n-1$ at most, and $P_n(x, y) = (ax+by+c)\pi_{n-1}(x, y)$, the identity states that

$$\int_C \rho(u, v)K_n(x, y, u, v)(au + bv + c)\pi_{n-1}(u, v)ds = (ax + by + c)\pi_{n-1}(x, y). \tag{2}$$

Since the points (x_ν, y_ν), $\nu = 1, 2, 3$, lie on the curve and at the same time on the line $ax+by+c=0$, the relation holds with the right-hand member equal to zero if any one of these pairs of values is substituted for (x, y). The functions $K_n(x_\nu, y_\nu, u, v)$ as polynomials in u and v are orthogonal to every polynomial of lower degree for weight $(au+bv+c)\rho(u, v)$.

The three polynomials $K_n(x_\nu, y_\nu, u, v)$ are linearly independent on C if $n \geqq 2$. For, as the polynomials $p_{kj}(u, v)$ are linearly independent on the curve, a relation of linear dependence of the form $\sum_\nu C_\nu K_n(x_\nu, y_\nu, u, v) = 0$ would imply corresponding relations of linear dependence among the quantities $p_{kj}(x_\nu, y_\nu)$ as coefficients in the K's. In particular, the first six of the equations $\sum_\nu C_\nu p_{kj}(x_\nu, y_\nu)$ would require that

$$\begin{aligned}
C_1 + C_2 + C_3 &= 0,\\
C_1x_1 + C_2x_2 + C_3x_3 &= 0,\\
C_1y_1 + C_2y_2 + C_3y_3 &= 0,\\
C_1x_1^2 + C_2x_2^2 + C_3x_3^2 &= 0,\\
C_1x_1y_1 + C_2x_2y_2 + C_3x_3y_3 &= 0,\\
C_1y_1^2 + C_2y_2^2 + C_3y_3^2 &= 0.
\end{aligned}$$

The first, second, and fourth of these equations are possible with C's not all zero only if two of the x's are equal, and the first, third, and

sixth only if two of the y's are equal. But this is inconsistent with the hypothesis that (x_1, y_1), (x_2, y_2), (x_3, y_3) are three distinct points on a straight line. Similarly, though more simply, any two of the polynomials $K_1(x_r, y_r, u, v)$ are linearly independent on C. The expressions $K_n(x_r, y_r, u, v)$ give the complete system of orthogonal polynomials corresponding to the weight function $(ax+by+c)\rho(x, y)$. If two of the points (x_r, y_r) are conjugate complex, real polynomials for the orthogonal system can be obtained by taking the real and pure imaginary parts of the corresponding K's.

Similar reasoning is clearly effective for a curve of the Nth degree with a straight line meeting it in N distinct points. For a treatment of certain cases of coincident points of intersection the discussion will be limited to the curve of the third degree.

Suppose now that the line $ax+by+c=0$, instead of meeting the curve in three distinct points, is tangent to it at an ordinary point (x_1, y_1), and intersects it at a point (x_2, y_2) distinct from (x_1, y_1). It is still assumed that the part of the curve on which $\rho \neq 0$ is all on one side of the line, that is, that $ax+by+c$ is of constant sign there. At least one of the coefficients a, b is different from zero; without loss of generality it may be supposed that $b \neq 0$, that is, that the line is not parallel to the y-axis. In the neighborhood of the point (x_1, y_1), y is a single-valued function of x on the curve, with a derivative whose value reduces at (x_1, y_1) to $(-a/b)$, the slope of the tangent line. In this neighborhood on C, by differentiation of (2),

$$\begin{aligned}\int_C \rho(u, v)K_{nx}(x, y, u, v)(au + bv + c)\pi_{n-1}(u, v)ds \\ &= \frac{d}{dx}\left[(ax + by + c)\pi_{n-1}(x, y)\right] \\ &= (ax + by + c)\frac{d}{dx}\pi_{n-1}(x, y) + (a + by')\pi_{n-1}(x, y),\end{aligned}$$

where $K_{nx}(x, y, u, v)$ means the total derivative of K_n with respect to x, equal to $\partial K_n/\partial x + y'\partial K_n/\partial y$. At (x_1, y_1) both $ax+by+c$ and $a+by'$ vanish. The function $K_{nx}(x_1, y_1, u, v)$, which is still a polynomial as regards its dependence on u and v, is orthogonal to $\pi_{n-1}(u, v)$ for weight $(au+bv+c)\rho(u, v)$.

There have been obtained then for $n \geqq 2$ three polynomials $K_n(x_1, y_1, u, v)$, $K_{nx}(x_1, y_1, u, v)$, $K_n(x_2, y_2, u, v)$, each possessing the property of orthogonality. These are once more linearly independent on the curve. Without effect on the form of K_n, and so without any impairment of generality, it may be assumed that the first six poly-

nomials in the orthogonal system are specifically those obtained by applying the Schmidt process to the monomials $1, x, y, x^2, xy, y^2$, taken in that order. Then $p_{11}(x, y)$ does not involve y, and $p_{21}(x, y)$ involves neither xy nor y^2. Suppose there were an identity of the form

$$C_1K_n(x_1, y_1, u, v) + C_2K_{n-1}(x_1, y_1, u, v) + C_3K_n(x_2, y_2, u, v) = 0$$

with C's not all zero. Then

$$C_1p_{kj}(x_1, y_1) + C_2p'_{kj}(x_1, y_1) + C_3p_{kj}(x_2, y_2) = 0$$

for each pair of indices (k, j) in which $k \leqq n$, the accent indicating total differentiation with respect to x. The equations for the first four pairs of indices may be simplified to the form

$$\begin{aligned} C_1 + C_3 &= 0, \\ C_1x_1 + C_2 + C_3x_2 &= 0, \\ C_1y_1 - C_2a/b + C_3y_2 &= 0, \\ C_1x_1^2 + 2C_2x_1 + C_3x_2^2 &= 0. \end{aligned}$$

The first, second, and fourth of these equations require that $x_1 = x_2$, which is impossible under the hypothesis that (x_1, y_1) and (x_2, y_2) are distinct and $b \neq 0$. It is immediately apparent that any two of the three K's are linearly independent for $n = 1$. So the K's give the complete orthogonal system for weight $(ax+by+c)\rho(x, y)$.

Finally, let the line meet the curve at a double point (x_1, y_1) at which there are distinct tangents with finite slopes, and at another point (x_2, y_2). It is assumed, as always, that $ax+by+c$ is of constant sign where ρ does not vanish. Let $K_1(u, v)$ and $K_2(u, v)$ be the total derivatives $K_{n-1}(x_1, y_1, u, v)$ calculated respectively for the two branches of the curve through (x_1, y_1); let the corresponding slopes be λ_1 and λ_2. Then

$$\int_C \rho(u, v)K_\nu(u, v)(au + bv + c)\pi_{n-1}(u, v)ds = (a + b\lambda_\nu)\pi_{n-1}(x_1, y_1), \qquad \nu = 1, 2.$$

It is seen that $(a+b\lambda_2)K_1(u, v) - (a+b\lambda_1)K_2(u, v)$ is orthogonal to $\pi_{n-1}(u, v)$ for the composite weight function. This polynomial and the two polynomials $K_n(x_1, y_1, u, v)$, $K_n(x_2, y_2, u, v)$ are found to be linearly independent on C for $n \geqq 2$, and once more a complete orthogonal system is obtained.

These illustrations will be allowed to suffice for the case of multiple intersections.

4. **Other cases.** The corresponding theory for a non-algebraic curve is complicated by the fact that the number of polynomials of the nth degree in the orthogonal system increases with n, but is compensatingly simplified by the observation that the representation corresponding to (2) holds for *all* values of x and y, so that (except for the assumption that the domain of orthogonality is all on one side of the line) the points where the straight line meets the curve are no longer a matter of special concern. If (x_r, y_r) are any $n+1$ distinct points on the line, independent polynomials of the nth degree orthogonal to every polynomial of lower degree with respect to the composite weight function are given by $K_n(x_r, y_r, u, v)$.

A similar conclusion holds for orthogonality on a two-dimensional region.

THE UNIVERSITY OF MINNESOTA

NOTE ON AN INEQUALITY OF STEINER[1]

T. RADÓ AND P. REICHELDERFER

Let Q denote the unit square $0 \leqq x, y \leqq 1$. If $f(x, y)$ be any function defined and continuous on Q, the relation $z=f(x, y)$ yields a continuous surface defined over Q. The Lebesgue area[2] of this surface will be denoted by $L(f)$. Let $z=f_1(x, y)$, $z=f_2(x, y)$ be two continuous surfaces defined over Q; then clearly $z=[f_1(x, y)+f_2(x, y)]/2$ is a continuous surface defined over Q. The *inequality of Steiner*[3] states that $L([f_1+f_2]/2) \leqq [L(f_1)+L(f_2)]/2$. McShane[4] obtained interesting and important results concerning the situation where the sign of equality holds in this relation. In this note we improve his results and, in a sense, give them a final form.

In order to emphasize and to clarify what is significant and interesting in the results of McShane and in our improvements thereon, we remind our reader of a few facts concerning the Lebesgue area.[5] Given a continuous surface $z=f(x, y)$ defined over Q; if $L(f)$ is finite then the partial derivatives f_x and f_y exist almost everywhere in Q, the integral $\iint_Q [1+f_x^2+f_y^2]^{1/2} dxdy$ exists, and the relation

[1] Presented to the Society, April 13, 1940.

[2] See S. Saks, *Theory of the Integral*, Warsaw and Lwów, 1937, chap. 5, for the facts used in this paper concerning the Lebesgue area.

[3] E. J. McShane, *On a certain inequality of Steiner*, Annals of Mathematics, (2), vol. 33 (1932), pp. 125–138.

[4] Loc. cit.[3]

[5] Cf.[2]

$\iint_Q [1+f_x^2+f_y^2]^{1/2}dxdy \leqq L(f)$ is valid. The sign of equality holds in this last relation if and only if $f(x, y)$ is absolutely continuous in the sense of Tonelli. Next, if a sequence of continuous functions $f_n(x, y)$ converges uniformly on Q to $f(x, y)$, then it is always true that $\liminf_{n\to\infty} L(f_n) \geqq L(f)$, but the sign of equality holds in this last relation only if the sequence $f_n(x, y)$ is chosen with extreme care.

Suppose $z=f_1(x, y)$, $z=f_2(x, y)$ are two continuous surfaces defined over Q such that $L([f_1+f_2]/2) = [L(f_1)+L(f_2)]/2$. If we assume, for the moment, that both $f_1(x, y)$ and $f_2(x, y)$ are absolutely continuous in the sense of Tonelli, then so is $[f_1(x, y)+f_2(x, y)]/2$, and the preceding equality is equivalent to

$$\iint_Q \left\{ \frac{[1+f_{1x}^2+f_{1y}^2]^{1/2}+[1+f_{2x}^2+f_{2y}^2]^{1/2}}{2} - \left[1+\left(\frac{f_{1x}+f_{2x}}{2}\right)^2+\left(\frac{f_{1y}+f_{2y}}{2}\right)^2\right]^{1/2} \right\} dxdy = 0. \tag{1}$$

Using the inequality of Schwartz one sees that the integrand in (1) is never negative; thus the integrand is equal to zero almost everywhere in Q, which implies that $f_{1x}=f_{2x}$, $f_{1y}=f_{2y}$ almost everywhere in Q—that is, $[(f_{1x}-f_{2x})^2+(f_{1y}-f_{2y})^2]^{1/2}=0$ almost everywhere in Q.[6]

Now suppose we discard the assumptions that f_1 and f_2 are absolutely continuous in the sense of Tonelli; then, of course, the preceding reasoning yields no result. Still McShane proves that $[(f_{1x}-f_{2x})^2+(f_{1y}-f_{2y})^2]^{1/2}=0$ almost everywhere in Q. Indeed, he proves a sharper result to the effect that, when $[L(f_1)+L(f_2)]/2 - L([f_1+f_2]/2)$ is small, then $[(f_{1x}-f_{2x})^2+(f_{1y}-f_{2y})^2]^{1/2}$ is small *in measure*, so to speak—quite exactly, he establishes a contrapositive statement:[7] *If the functions $f_1(x, y)$, $f_2(x, y)$ are defined and continuous on a square Q and there exist positive numbers G, ϵ, δ such that*

(1) $L(f_1) \leqq G$, $L(f_2) \leqq G$;

(2) $[(f_{1x}-f_{2x})^2+(f_{1y}-f_{2y})^2]^{1/2} \geqq \epsilon$ *on a set of measure at least equal to δ;*

then

$$L([f_1+f_2]/2) \leqq [L(f_1)+L(f_2)]/2 - \mu(G, \epsilon, \delta),$$

where μ is a positive number whose value depends only on G, ϵ, and δ.

We improve this result by replacing smallness with respect to

[6] From this it follows (loc. cit.,[3] Theorem I) that f_1-f_2 is constant on Q, but we shall not be concerned with this fact.

[7] Loc. cit.,[3] Theorem IV.

measure by smallness with respect to *exponent*. First, we show that, *if $f_1(x, y)$ and $f_2(x, y)$ are two continuous functions defined on the square Q for which $L(f_1)$ and $L(f_2)$ are both finite, then*

$$\left\{\iint_Q [(f_{1x}-f_{2x})^2+(f_{1y}-f_{2y})^2]^{1/4}dxdy\right\}^4 \leqq [L(f_1)+L(f_2)]^3\left[\frac{L(f_1)+L(f_2)}{2}-L\left(\frac{f_1+f_2}{2}\right)\right]. \tag{2}$$

Since the Hölder inequality reveals

$$\left\{\iint_Q [(f_{1x}-f_{2x})^2+(f_{1y}-f_{2y})^2]^{\lambda/2}dxdy\right\}^{2/\lambda}$$

to be a non-decreasing function of λ, it follows that, for every exponent λ satisfying $0<\lambda\leqq 1/2$, it is true that

$$\left\{\iint_Q [(f_{1x}-f_{2x})^2+(f_{1y}-f_{2y})^2]^{\lambda/2}dxdy\right\}^{2/\lambda} \leqq [L(f_1)+L(f_2)]^3\left[\frac{L(f_1)+L(f_2)}{2}-L\left(\frac{f_1+f_2}{2}\right)\right]. \tag{3}$$

Next, we prove that, *under the same assumptions, it follows that*

$$\left\{\iint_Q [(f_{1x}-f_{2x})^2+(f_{1y}-f_{2y})^2]^{\lambda/2}dxdy\right\}^{2/\lambda} \leqq [L(f_1)+L(f_2)]^{(1+\lambda)/\lambda}\left[\frac{L(f_1)+L(f_2)}{2}-L\left(\frac{f_1+f_2}{2}\right)\right]^{(1-\lambda)/\lambda} \tag{4}$$

for every exponent λ satisfying $1/2\leqq\lambda\leqq 1$. The reader will observe that inequality (4) reduces to (2) for $\lambda=1/2$.

Let, now, $f_n(x, y)$ be a sequence of continuous functions converging uniformly on Q to $f(x, y)$; assume that $L(f_n)$ and $L(f)$ are finite, and that $L(f_n)$ converges to $L(f)$. Under these hypotheses McShane proves that $[(f_{nx}-f_x)^2+(f_{ny}-f_y)^2]^{1/2}$ converges to zero *in measure*.[8] We improve this result by showing that *under these hypotheses it is true that*

$$\lim_{n\to\infty}\iint_Q [(f_{nx}-f_x)^2+(f_{ny}-f_y)^2]^{\lambda/2}dxdy=0 \tag{5}$$

for every exponent λ satisfying $0<\lambda<1$. Since McShane has shown

[8] Loc. cit.,[2] Theorem V.

that the relation (5) is not generally true for the exponent $\lambda = 1$,[9] our result is the best obtainable under the given hypotheses.

In fact, we shall give two proofs for relation (5). One proof will reveal this result as an immediate consequence of inequalities (3) and (4); the other will depend upon a general lemma concerning the Lebesgue integral.

We proceed presently to the proof of inequality (2). Given two continuous functions $f_1(x, y)$, $f_2(x, y)$ defined on Q for which $L(f_1)$, $L(f_2)$ are both finite; then $f_{1x}, f_{1y}; f_{2x}, f_{2y}$ exist almost everywhere in Q. Set

$$u = [(f_{1x} - f_{2x})^2 + (f_{1y} - f_{2y})^2]^{1/2}; \quad v = 1 + f_{1x}f_{2x} + f_{1y}f_{2y};$$
$$w_1 = [1 + f_{1x}^2 + f_{1y}^2]^{1/2}; \quad w_2 = [1 + f_{2x}^2 + f_{2y}^2]^{1/2};$$
$$w = [1 + ([f_{1x} + f_{2x}]/2)^2 + ([f_{1y} + f_{2y}]/2)^2]^{1/2};$$
$$\alpha = (w_1 + w_2)/2 - w; \quad \beta = (w_1 + w_2)/2 + w;$$
$$\gamma = w_1w_2 - v; \quad \delta = w_1w_2 + v.$$

The reader will easily verify the following identities:

$$\alpha + \beta = w_1 + w_2, \quad \alpha \cdot \beta = \gamma/2, \tag{6}$$

$$\gamma + \delta = 2w_1w_2, \quad \gamma \cdot \delta = u^2 + (f_{1x}f_{2y} - f_{2x}f_{1y})^2. \tag{7}$$

Now (7) implies that $\gamma \geqq 0$, $\delta \geqq 0$, whence it follows that $\alpha \cdot \beta \geqq 0$, and that

$$\delta \leqq 2w_1w_2 \leqq (w_1 + w_2)^2/2.$$

In view of these relations (6) implies that $\alpha \geqq 0, \beta \geqq 0$; hence $\beta \leqq w_1 + w_2$. So the reader sees that

$$u^2 \leqq \gamma\delta = 2\alpha\beta\delta \leqq (w_1 + w_2)^3\{(w_1 + w_2)/2 - w\}. \tag{8}$$

As we observed above, it is true that

$$\iint_Q w_i\,dx\,dy \leqq L(f_i), \qquad i = 1, 2. \tag{9}$$

A reasoning of McShane[10] shows that

$$\iint_Q \{(w_1 + w_2)/2 - w\}\,dx\,dy \leqq [L(f_1) + L(f_2)]/2 - L([f_1 + f_2]/2). \tag{10}$$

Using the Hölder inequality the reader will speedily verify that inequality (2) follows from (8), (9), and (10).

[9] Loc. cit.,[3] Theorem VI.

[10] Loc. cit.,[3] p. 129.

We proceed next to establish inequality (4). It follows by the triangle inequality that $u \leqq w_1+w_2$; hence we have from (9),

$$\iint_Q [(f_{1x}-f_{2x})^2+(f_{1y}-f_{2y})^2]^{1/2}dxdy \leqq L(f_1)+L(f_2). \tag{11}$$

Suppose that λ_2 is a positive number and $F(x, y)$ is a non-negative function such that F^{λ_2} is summable on Q. If λ_1 be any positive number less than λ_2, it follows by the Hölder inequality that F^{λ_1} is also summable on Q. Now every number λ satisfying $\lambda_1 \leqq \lambda \leqq \lambda_2$ may be written in the form $\lambda=\alpha_1\lambda_1+\alpha_2\lambda_2$, where $0 \leqq \alpha_1, \alpha_2 \leqq 1$, $\alpha_1+\alpha_2=1$. From the Hölder inequality we have

$$\iint_Q F^{\lambda}dxdy \leqq \left[\iint_Q F^{\lambda_1}dxdy\right]^{\alpha_1}\left[\iint_Q F^{\lambda_2}dxdy\right]^{\alpha_2}. \tag{12}$$

The reader will verify that (4) follows from (2), (11), and (12) with $F=u, \lambda_1=1/2, \lambda_2=1$.

Now let $f_n(x, y)$ be a sequence of continuous functions converging uniformly on Q to $f(x, y)$; assume that $L(f_n)$ and $L(f)$ are finite, and that $L(f_n)$ converges to $L(f)$. We first show that

$$\lim_{n\to\infty}\left[\frac{L(f_n)+L(f)}{2}-L\left(\frac{f_n+f}{2}\right)\right]=0. \tag{13}$$

Since $(f_n+f)/2$ converges uniformly to f on Q, it follows that

$$\liminf_{n\to\infty} L\left(\frac{f_n+f}{2}\right) \geqq L(f). \tag{14}$$

But $\lim_{n\to\infty}[L(f_n)+L(f)]/2=L(f)$, and by the inequality of Steiner it follows that $L([f_n+f]/2) \leqq [L(f_n)+L(f)]/2$. Thus

$$\limsup_{n\to\infty} L\left(\frac{f_n+f}{2}\right) \leqq L(f). \tag{15}$$

From (14) and (15), the assertion in (13) follows.

Now the inequality (3) shows that, for $0<\lambda \leqq 1/2$, it is true that

$$\begin{aligned}\iint_Q [(f_{nx}-f_x)^2+(f_{ny}-f_y)^2]^{\lambda/2}dxdy \\ \leqq [L(f_n)+L(f)]^{2\lambda/2}\left[\frac{L(f_n)+L(f)}{2}-L\left(\frac{f_n+f}{2}\right)\right]^{\lambda/2},\end{aligned} \tag{16}$$

while inequality (4) shows that, for $1/2 \leqq \lambda \leqq 1$, it is true that

$$\iint_Q [(f_{nx} - f_x)^2 + (f_{ny} - f_y)^2]^{\lambda/2} dxdy$$

$$\leqq [L(f_n) + L(f)]^{(1+\lambda)/2} \left[\frac{L(f_n) + L(f)}{2} - L\left(\frac{f_n + f}{2}\right)\right]^{(1-\lambda)/2} \tag{17}$$

From (13), (16), (17) the relation (5) clearly follows.

In order to give a second proof of relation (5) we use the following lemma.[11] *Let $F_n(x, y)$ be a sequence of non-negative functions defined on Q for which there exists a positive number λ_0 such that* (1) *the functions F_n converge to zero in measure*; (2) *the functions $F_n^{\lambda_0}$ are summable on Q*; (3) *the integrals $\iint_Q F_n^{\lambda_0} dxdy$ are uniformly bounded by a positive number M. Then, for every positive number λ less than λ_0, it is true that* $\lim_{n\to\infty} \iint_Q F_n^{\lambda} dxdy = 0$.

PROOF. Given $\epsilon > 0$, since F_n converges to zero in measure there exists a positive integer $n(\epsilon)$ such that, for every $n > n(\epsilon)$ the set $E_n(\epsilon)$ of points (x, y) in Q for which $F_n(x, y) \geqq \epsilon^{1/\lambda}$ is of measure less than $\epsilon^{\lambda_0/(\lambda_0-\lambda)} \cdot M^{-\lambda/(\lambda_0-\lambda)}$. Now obviously

$$\iint_Q F_n^{\lambda} dxdy = \iint_{Q-E_n(\epsilon)} F_n^{\lambda} dxdy + \iint_{E_n(\epsilon)} F_n^{\lambda} dxdy. \tag{18}$$

On $Q - E_n(\epsilon)$ it is clear that $F_n^{\lambda}(x, y) \leqq \epsilon$; hence

$$\iint_{Q-E_n(\epsilon)} F_n^{\lambda} dxdy \leqq \epsilon. \tag{19}$$

Using the Hölder inequality, we have, for $n > n(\epsilon)$,

$$\iint_{E_n(\epsilon)} F_n^{\lambda} dxdy$$

$$\leqq \left[\iint_{E_n(\epsilon)} F_n^{\lambda_0} dxdy\right]^{\lambda/\lambda_0} \left[\iint_{E_n(\epsilon)} dxdy\right]^{(\lambda_0-\lambda)/\lambda_0} < \epsilon. \tag{20}$$

Since ϵ is arbitrary, the lemma follows immediately from (18), (19), (20).

Now let $f_n(x, y)$ be a sequence of continuous functions converging uniformly on Q to a function $f(x, y)$; assume that $L(f_n)$ and $L(f)$ are finite, and that $\lim_{n\to\infty} L(f_n) = L(f)$. Consider the sequence of functions $F_n(x, y) = [(f_{nx} - f_x)^2 + (f_{ny} - f_y)^2]^{1/2}$. In view of the results of McShane

[11] This lemma is an immediate corollary of certain important results of F. Riesz, *Untersuchungen über Systeme integrierbarer Funktionen*, Mathematische Annalen, vol. 69 (1910), pp. 449–497. We give a direct proof for the convenience of the reader.

stated above, it is clear that this sequence of functions satisfies the three conditions of the preceding lemma with $\lambda_0=1$. Thus, for every exponent λ satisfying $0<\lambda<1$ it is true that

$$\lim_{n\to\infty}\iint_Q [(f_{nx}-f_x)^2+(f_{ny}-f_y)^2]^{\lambda/2}dxdy=0.$$

So we have a second proof of our result (5).

THE OHIO STATE UNIVERSITY

AN ADDITIONAL CRITERION FOR THE FIRST CASE OF FERMAT'S LAST THEOREM[1]

BARKLEY ROSSER

In an earlier paper[2] it was shown that if p is an odd prime and

$$a^p + b^p + c^p = 0$$

has a solution in integers prime to p, then

$$m^{p-1} \equiv 1 \pmod{p^2}$$

for each prime $m \leqq 41$. In this paper the result is extended to $m \leqq 43$.

We will use the notations and conventions of I throughout, and a reference to a numbered equation will refer to the equation of that number in I. With p assumed to be an odd prime such that (1) has a solution in integers prime to p, we assume that a t exists such that the values of (2) satisfy (4), (5), and (6) with $m=43$. Put $g(x)=f(x)f(-x)$ and

$$h(x) = (x^{42} - 1)/(x^6 - 1).$$

Then $g(x)$ divides $h(x)$, and $g(x)$ can be completely factored modulo p.

Case 1. Assume that a root of $g(x)$ is a root of

$$h(x)/(x^{12} + x^{10} + x^8 + x^6 + x^4 + x^2 + 1).$$

Then this root belongs to either the exponent 21 or the exponent 42 modulo p. Hence $p \equiv 1 \pmod{42}$. So there is an ω such that

$$\omega^2 + \omega + 1 \equiv 0.$$

Then $g(x)$, $g(\omega x)$, and $g(\omega^2 x)$ all divide $h(x)$. Moreover, the only cases in which two of $g(x)$, $g(\omega x)$, and $g(\omega^2 x)$ have a common factor are

I. $a^6+1 \equiv 0$,

II. $a^6+a^3+3a^2+3a+1 \equiv 0$,

III. $a^6-a^3-3a^2-3a-1 \equiv 0$,

or cases derived from these by replacing a by one of the other roots of $f(x)$. So if we show that $h(x)$ has no factor in common with any of x^6+1, $x^6+x^3+3x^2+3x+1$, or $x^6-x^3-3x^2-3x-1$, then we can conclude that $g(x)g(\omega x)g(\omega^2 x)$ must divide $h(x)$.

Clearly $h(x)$ has no factor in common with x^6+1.

---

[1] Presented to the Society, April 27, 1940.

[2] *A new lower bound for the exponent in the first case of Fermat's last theorem*, this Bulletin, vol. 46 (1940), pp. 299–304. This paper will be referred to as I.

Suppose $h(x)$ has a factor in common with $x^6+x^3+3x^2+3x+1$. This latter has the factors x^2+x+1 and x^4-x^3+2x+1. The first has no factor in common with $h(x)$, since it divides x^6-1, which has no factor in common with $h(x)$. To test the second, we try it successively with each of the four factors of $h(x)$, getting the eliminants

$$13\cdot 19^2\cdot 127\cdot 163^2,\ 5\cdot 36913,\ 2\cdot 127,\ 5\cdot 7.$$

Suppose $h(x)$ has a factor in common with $x^6-x^3-3x^2-3x-1$. This latter has the factors x^2-x-1 and $x^4+x^3+2x^2+2x+1$. The first has no factor in common with $h(x)$ by Lemma 3 of I. Trying the second factor successively with each of the four factors of $h(x)$, we get the eliminants

$$7^3\cdot 43,\ 2^2\cdot 7\cdot 13\cdot 43,\ 7,\ 43.$$

So $g(x)g(\omega x)g(\omega^2 x)$ must divide $h(x)$. Since both are of degree 36, they must be equal. Putting $b=c+5$ and equating coefficients, we get

$$A+1=2c^3+3c^2-24c+13\equiv 1,$$

$$B+1=c^6+12c^5+42c^4+18c^3-9c^2-222c+173\equiv 1,$$

$$C+1=-2c^6+12c^5+171c^4+132c^3-666c^2+132c+201\equiv 1.$$

Dividing $16B$ and $8C$ by A, we get the remainder

$$43D=43(99c^2+192c-116)\equiv 0$$

from each. Then

$$2cE=29A+3D=2c(29c^2+192c-60)\equiv 0.$$

As $c\equiv 0$ would give $A\equiv 12\equiv 0$, we have

$$28cF=15D-29E=28c(23c-96)\equiv 0,$$

$$29cG=8E-5F=29c(8c-49)\equiv 0,$$

$$8F-23G=359\equiv 0.$$

Case 2. Assume that no root of $g(x)$ is a root of

$$h(x)/(x^{12}+x^{10}+x^8+x^6+x^4+x^2+1).$$

Then, since $g(x)$ divides $h(x)$ and is of degree 12,

$$g(x)\equiv x^{12}+x^{10}+x^8+x^6+x^4+x^2+1.$$

So $2c+1\equiv 1$ and $c^2+5\equiv 1$.

PRINCETON UNIVERSITY

SOME NEW PROPERTIES OF TRANSFINITE ORDINALS[1]

SEYMOUR SHERMAN

1. **Introduction.** Sierpiński[2] has stated the following theorem on the Cantor normal form:

Every ordinal $A>0$ can be represented in the form

$$A = \sum_{i=0}^{n-1} \omega^{\alpha_i} a_i,$$

where

$$\omega > n,\ a_0,\ a_1,\ a_2,\ \cdots,\ a_{n-1} > 0,$$

$$A \geqq \alpha_0 > \alpha_1 > \alpha_2 > \cdots > \alpha_{n-1} \geqq 0.$$

Henceforth in this paper, when an ordinal is written in summation form, it will be assumed that the summation satisfies the above-stated requirements.

Elsewhere Sierpiński[3] has proved:

The set of all "divisors on the left" of an ordinal number is closed.[4]

One of the purposes of this paper is to analyze further, by means of the Cantor normal form, the left and right factors of transfinite ordinals. In addition the Cantor normal form will be used to prove that for ordinals A, B, and C

$$(A + B)C \leqq AC + BC.$$

We can also specify the necessary and sufficient condition that the equality hold. F. Siecza[5] has used the Cantor normal form to discuss

---

[1] Presented to the Society, December 27, 1939. The author is indebted to Professor W. A. Hurwitz for his indispensable advice in the preparation of this paper.

[2] W. Sierpiński, *Leçons sur les Nombres Transfinis*, Paris, 1928, p. 202; hereafter referred to as Sierpiński (I). The theorem was first stated for ordinals of the first and second ordinal class by G. Cantor, *Beiträge zur Begründung der transfiniten Mengenlehre*, Mathematische Annalen, vol. 49 (1897), p. 237, and for all ordinals by G. Hessenberg, *Grundbegriffe der Mengenlehre*, Abhandlungen der Fries'schen Schule, New Series 1.4, Göttingen, 1906, p. 587.

[3] W. Sierpiński, *A property of ordinal numbers*, Bulletin of the Calcutta Mathematical Society, vol. 20 (1930), pp. 21–22; hereafter referred to as Sierpiński (II).

[4] A set of ordinal numbers is said to be closed if it contains the least upper bound of each subset.

[5] F. Siecza, *Sur l'unicité de la décomposition de nombres ordinaux en facteurs irreductibles*, Fundamenta Mathematicae, vol. 5 (1924), pp. 172–175.

factorization of transfinite ordinals, but his problem was different from the one considered here.

2. **Left and right factors.** We shall now consider left factors.

2.1. THEOREM. *A necessary and sufficient condition that x be a left factor of $A=\sum_{i=0}^{n-1}\omega^{\alpha_i}a_i$ is that either $0<x<\omega^{\alpha_{n-1}}$ or $x=\omega^{\alpha_j}p_j+\sum_{i=j+1}^{n-1}\omega^{\alpha_i}a_i$, where p_j is a factor of a_j and $0\leqq j\leqq n-1$.*

If $0<x<\omega^{\alpha_{n-1}}$, let $x=\sum_{i=0}^{m-1}\omega^{\beta_i}b_i$. Thus $\beta_0<\alpha_{n-1}$ and $-\beta_0+\alpha_j$, $0\leqq j\leqq n-1$, exists.[6] We have

$$\left\{\sum_{i=0}^{m-1}\omega^{\beta_i}b_i\right\}\left\{\sum_{i=0}^{n-1}\omega^{-\beta_0+\alpha_i}\right\}=\sum_{i=0}^{n-1}\omega^{\alpha_i}a_i,$$

and hence x is a left factor of A.

If $x=\omega^{\alpha_j}p_j+\sum_{i=j+1}^{n-1}\omega^{\alpha_i}a_i$, where $a_j=p_j\lambda_j$, then

$$\left\{\omega^{\alpha_j}p_j+\sum_{i=j+1}^{n-1}\omega^{\alpha_i}a_i\right\}\left\{\sum_{i=0}^{j-1}(\omega^{-\alpha_j+\alpha_i})a_i+\lambda_j\right\}$$

$$=\sum_{i=0}^{j-1}(\omega^{\alpha_j-\alpha_j+\alpha_i})a_i+\omega^{\alpha_j}p_j\lambda_j+\sum_{i=j+1}^{n-1}\omega^{\alpha_i}a_i$$

$$=\sum_{i=0}^{j-1}\omega^{\alpha_i}a_i+\omega^{\alpha_j}a_j+\sum_{i=j+1}^{n-1}\omega^{\alpha_i}a_i=\sum_{i=0}^{n-1}\omega^{\alpha_i}a_i$$

and x is a left factor of A.

Thus we have proved the sufficiency of the condition.

Suppose that $x=\sum_{i=0}^{m-1}\omega^{\beta_i}b_i$ is a left factor of A and that $x\geqq\omega^{\alpha_{n-1}}$. Then

$$A=\left(\sum_{i=0}^{m-1}\omega^{\beta_i}b_i\right)\left(\sum_{i=0}^{l-1}\omega^{\gamma_i}c_i\right).$$

If $\gamma_{l-1}=0$ then

$$\sum_{i=0}^{n-1}\omega^{\alpha_i}a_i=\left(\sum_{i=0}^{m-1}\omega^{\beta_i}b_i\right)\left(\sum_{i=0}^{l-1}\omega^{\gamma_i}c_i\right)$$

$$=\sum_{i=0}^{l-2}(\omega^{\beta_0+\gamma_i})c_i+\omega^{\beta_0}b_0c_{l-1}+\sum_{i=1}^{m-1}\omega^{\beta_i}b_i,$$

where the last expression is a Cantor normal form. Comparing coefficients in the right and left ends of the series of equations, we see that

[6] See F. Hausdorff, *Mengenlehre*, 1935, p. 63.

for some j, $0 \leqq j \leqq n-1$, $\beta_0 = \alpha_j$, b_0 is a factor of a_j, and $\sum_{i=1}^{m-1} \omega^{\beta_i} b_i = \sum_{i=j+1}^{n-1} \omega^{\alpha_i} a_i$.

If $\gamma_{l-1} > 0$, then

$$\sum_{i=0}^{n-1} \omega^{\alpha_i} a_i = \left(\sum_{i=0}^{m-1} \omega^{\beta_i} b_i \right) \left(\sum_{i=0}^{l-1} \omega^{\gamma_i} c_i \right) = \sum_{i=0}^{l-1} (\omega^{\beta_0 + \gamma_i}) c_i.$$

But since $x \geqq \omega^{\alpha_{n-1}}$, we have $\beta_0 \geqq \alpha_{n-1}$, $\beta_0 + \gamma_{l-1} > \alpha_{n-1}$, and thus the left and right ends of series of equations could not be equal. From this contradiction $\gamma_{l-1} = 0$ and by the preceding argument our condition is necessary.

Since in $\sum_{i=0}^{n-1} \omega^{\alpha_i} a_i$ there is only a finite set of coefficients and since each of these coefficients, being finite, has only a finite set of factors, there is but a finite set of left factors of the form $\omega^{\alpha_j} p_j + \sum_{i=j+1}^{n-1} \omega^{\alpha_i} a_i$, p_j a factor of a_j; that is, there is only a finite set of left factors greater than or equal to $\omega^{\alpha_{n-1}}$. Hence we have proved:

2.2. COROLLARY. *The set of left factors of* $A = \sum_{i=0}^{n-1} \omega^{\alpha} a_i$ *consists of the set of all positive ordinals less than* $\omega^{\alpha_{n-1}}$ *and a finite set of ordinals not less than* $\omega^{\alpha_{n-1}}$.

2.3. COROLLARY. *If* $A = \sum_{i=0}^{n-1} \omega^{\alpha_i} a_i$, *the left factors of* A, *in order of magnitude, form a normally ordered set whose ordinal number is* $-1 + \omega^{\alpha_{n-1}} + \sum_{i=0}^{n-1} \tau(a_i)$, *where* $\tau(a_i)$ *is the number of factors of* a_i.

2.31. COROLLARY. *The ordinal number of the normally ordered set of left factors of a positive ordinal is either a finite ordinal or the sum of an additively indecomposable*[7] *ordinal and a finite ordinal.*

We can now obtain the previously stated result of Sierpiński,[8] who makes no use of the Cantor normal form:

2.4. COROLLARY. *The left factors of* A *form a closed set.*

The following corollary has been proved by Hausdorff[9] and Bennett[10] by means of the euclidean algorithm.

2.5. COROLLARY. *Any two ordinals greater than* 0 *have a greatest common left factor.*

We shall prove this by means of the preceding corollary. Let $A > 0$

[7] An ordinal $\alpha > 0$ is said to be additively indecomposable when there do not exist ordinals β, γ such that $\beta < \alpha$, $\gamma < \alpha$, and $\alpha = \beta + \gamma$. See Sierpiński (I), p. 178.

[8] Sierpiński (II).

[9] F. Hausdorff, *Grundzüge der Mengenlehre*, Leipzig, 1914.

[10] A. A. Bennett, *Some arithmetic operations with transfinite ordinals*, American Mathematical Monthly, vol. 28 (1921), pp. 427–430.

have a set of left factors, $\mathfrak{A}$, and $B>0$ have a set of left factors $\mathfrak{B}$. $\mathfrak{AB}$, the intersection of $\mathfrak{A}$ and $\mathfrak{B}$, has at least one element, 1. The least upper bound, $\mathfrak{c}$, of $\mathfrak{AB}$ is a member of $\mathfrak{A}$ and a member of $\mathfrak{B}$, since both $\mathfrak{A}$ and $\mathfrak{B}$ are closed. Hence $\mathfrak{c}$ is a member of $\mathfrak{AB}$; that is, $\mathfrak{AB}$ has a maximum element. Thus $\mathfrak{c}$ is the greatest common left factor of A and B.

The development in case of right factors is somewhat analogous. The following can easily be seen to be true.

2.6. THEOREM. *If* $A=\sum_{i=0}^{n-1}\omega^{\alpha_i}a_i$, *then a necessary and sufficient condition that* x *be a right factor of* A *is that either* $x=\sum_{i=0}^{n-1}(\omega^{-\beta_0+\alpha_i})a_i$, *where* $\beta_0<\alpha_{n-1}$, *or* $x=\sum_{i=0}^{j-1}(\omega^{-\alpha_j+\alpha_i})a_i+r_j$, *where* r_j *is a factor of* a_j.

Since each exponent of the Cantor normal form has a finite set of residues and each coefficient has only a finite set of factors, we can state the following:[11]

2.7. THEOREM. *The set of all right factors of any positive ordinal is finite.*

Sierpiński has proved some of these theorems by using merely the properties of additively indecomposable ordinals. The use of the Cantor normal form not only yields more extensive results in the case of left factors, but also provides a unified approach to other problems in the theory of transfinite ordinals.

3. **The distributive laws.** Among transfinite ordinals the distributive law is always valid for one order of multiplication

$$C(A+B)=CA+CB,$$

but not always for the other order of multiplication

$$(A+B)C \quad ? \quad AC+BC.$$

We shall prove:

3.1. THEOREM. $(A+B)C\leqq AC+BC$.

If either A, B, or C is zero, then our theorem holds. Consider $A>0$, $B>0$, and $C>0$, where

$$A=\sum_{i=0}^{n-1}\omega^{\alpha_i}a_i=\omega^{\alpha_0}a_0+A', \qquad B=\sum_{i=0}^{m-1}\omega^{\beta_i}b_i=\omega^{\beta_0}b_0+B',$$

$$C=\sum_{i=0}^{l-1}\omega^{\gamma_i}c_i.$$

[11] Sierpiński (I).

C is either finite, transfinite and limiting, or transfinite and non-limiting.

Suppose C is finite and $0<C<\omega$. If $\alpha_0<\beta_0$, then

$$A + B = A + \omega^{\beta_0}b_0 + B' = B, \qquad (A + B)C = BC = AC + BC.$$

If $\alpha_0=\beta_0$, then

$$A + B = \omega^{\beta_0}(a_0 + b_0) + B'$$

and

$$\begin{aligned} AC + BC &= \omega^{\alpha_0}a_0C + A' + \omega^{\beta_0}b_0C + B' \\ &= \omega^{\alpha_0}(a_0 + b_0)C + B' = (A + B)C. \end{aligned}$$

If $\alpha_0>\beta_0$, then

$$A + B = \omega^{\alpha_0}a_0 + A' + \omega^{\beta_0}b_0 + B',$$

and

$$(A + B)C = \omega^{\alpha_0}a_0C + A' + \omega^{\beta_0}b_0 + B' = AC + B \leqq AC + BC.$$

Thus if $0<C<\omega$, then

$$(A + B)C \leqq AC + BC.$$

Suppose C is transfinite and limiting, then $\gamma_{l-1}>0$. Let $J=\max(A, B)$; then

$$(A + B)C \leqq (J + J)C = (J2)C = J(2C) = JC \leqq AC + BC.$$

Suppose C is transfinite and non-limiting, then $\gamma_{l-1}=0$, and if we take

$$C_0 = \sum_{i=0}^{l-2} \omega^{\gamma_i}c_i, \qquad C_1 = C_{l-1},$$

then

$$C_0 + C_1 = C.$$

If $\alpha_0 \geqq \beta_0$, then

$$\begin{aligned} (A + B)C &= (A + B)(C_0 + C_1) = (A + B)C_0 + (A + B)C_1 \\ &\leqq AC_0 + AC_1 + BC_1 < AC + BC. \end{aligned}$$

If $\alpha_0<\beta_0$, then

$$(A + B)C \leqq BC \leqq AC + BC.$$

Thus if $\gamma_{l-1}=0$, then

$$(A + B)C \leqq AC + BC,$$

and our theorem is proved.

We can also prove the following:

3.2. COROLLARY. *A necessary and sufficient condition that*

$$(A + B)C = AC + BC$$

for positive A, B, and C is that either $C=1$, *or* $1<C<\omega$ *and* $\alpha_0 \leqq \beta_0$, *or* $\omega \leqq C$ *and* $\alpha_0+\gamma_0<\beta_0+\gamma_0$.

This corollary follows quite easily from the reasoning found in the preceding section.

CORNELL UNIVERSITY

THE DECOMPOSITION THEOREM FOR ABELIAN GROUPS[1]

JOEL BRENNER

Let G be an abelian group such that $p^k g = \mathbf{0}$ for all $g \in G$, p prime, k fixed. We prove *G has a basis*, that is, a set of elements such that each $g \in G$ is uniquely expressible as a linear combination of elements of the set.[2]

THEOREM. *There exists an ascending chain of sets* B_i, $0 \leqq i \leqq k$, *of elements of G with the properties*:

(i) *Every element in* B_i *is of order greater than* p^{k-i}.

(ii) *The elements in* B_i *are completely linearly independent.*

(iii) *If the order of the element g in G is greater than* p^{k-i}, *then there exists a (unique) linear combination z of elements of* B_i *such that the order of* $g-z$ *is at most* p^{k-i}.

Since we may choose as B_0 the vacuous set, we may assume that the sets $B_0, \cdots, B_s$ have already been constructed in such a way as to meet the requirements (i) to (iii). In order to construct B_{s+1} we adjoin to B_s any greatest subset C of G with the following properties.

(a) All the elements in C are of order p^{k-s}.

(b) The join B_{s+1} of the sets B_s and C is an independent set.

---

[1] Presented to the Society, April 6, 1940.

[2] Unique in that the number of nonzero terms in an expression for g is unique and only the arrangement but not the respective values of the nonzero terms may differ in two expressions for g.

The set B_{s+1} satisfies (i) and (ii). In order to prove (iii) let h be any element in G whose order is at least p^{k-s}. If firstly the order of h is exactly p^{k-s}, then it follows from the conditions (a) and (b) that there exists an integer m and a linear combination $y=\sum a_j y_j$ of elements in B_{s+1} so that $0 \neq mh = y$. Assume without loss of generality that m is a power of $p < p^{k-s}$. Then $(p^{k-s}/m)mh = 0 = \sum a_j(p^{k-s}/m)y_j$. By (ii), p^{k-s} divides $a_j(p^{k-s}/m)$, so that a_j/m is an integer (all j). Therefore the order of $h - \sum(a_j/m)y_j$ is $m < p^{k-s}$. If secondly the order of h is greater than p^{k-s}, then there exists a linear combination z of elements in B_s such that the order of $h-z$ is at most p^{k-s}, and thus (as above) $y=\sum c_j y_j$ can be found with y_j in B_{s+1} for which $h-z-y$ has order $< p^{k-s}$. We have shown[3] that B_k is a basis of G.

UNIVERSITY OF CALIFORNIA

[3] The following contain proofs for finite groups: Andreas Speiser, *Theorie der Gruppen von endlicher Ordnung*, 3d edition, Berlin, Julius Springer, 1937, p. 46–49. R. Remak, *Über die Zerlegung der kommutativen Gruppen in zyklische teilerfremde Faktoren*, Journal für die reine und angewandte Mathematik, vol. 141. L. C. Mathewson, *A simple proof of a theorem of Kronecher*, ibid., vol. 161 (1929), p. 255. A. Korselt, ibid. N. Tschebotaroew, *Bewies der Existenz einer Basis bei Abelschen Gruppen von endlicher Ordnung*, Kasan University, Fizzico-Matematico Obchestvo Izviestiia, vol. 4 (1929–1930). Third line after δ) in the proof read $\omega = \omega_s r$ for $\omega = \omega_s = r$. In the fourth paragraph following, read $\bar{\omega}_s = q\omega_s + t$, $0 \leq t < \omega$, $A_s \cdots t \cdots$. The following contain proofs for infinite groups: H. Pruefer, *Untersuchungen ueber die Zerlegbarkeit*, Mathematische Zeitschrift, vol. 17 (1923), p. 53; R. Baer, Compositio Mathematica, vol. 1 (1934), p. 274.

ON THE CONVEX SOLUTION OF A CERTAIN FUNCTIONAL EQUATION[1]

H. P. THIELMAN

The object of this paper is to give a generalization of a result obtained in a recent article by Mayer.[2] His result is a particular case (for $a=1$, $p=1$, $K=0$) of the following theorem.

THEOREM 1. *The only function which is convex for all $x>K\geqq 0$, and satisfies the functional equation*

$$1/f(x+a) = x^p f(x), \qquad x>0,\ a>0,\ p>0, \tag{1}$$

is

$$F(x) = \left[\frac{\Gamma(x/2a)}{(2a)^{1/2}\Gamma((x+a)/2a)}\right]^p. \tag{2}$$

The following proof follows closely that given by Mayer in the paper to which we referred above. All real functions occurring in this paper will be assumed to be defined for all positive values of the independent variable unless otherwise specified. A function $f(x)$ is said to be convex for all $x>K$ if for every pair of values of x, say x_1 and x_2, for which $f(x)$ is defined, and $x_1>K$, $x_2>K$,

$$f\left(\frac{x_1+x_2}{2}\right) \leqq \tfrac{1}{2}[f(x_1)+f(x_2)]. \tag{3}$$

That $f(x)$ satisfies the functional equation (1) follows from the fundamental equation for the Gamma function, $\Gamma(x+1)=x\Gamma(x)$.

We shall now show that $F(x)$ is convex. From (2) it follows[3] that

$$\frac{F'(x)}{F(x)} = -ap\sum_{v=0}^{\infty}\frac{1}{(2av+x)(2av+a+x)}.$$

Since $a>0$, $p>0$, it follows that $F'(x)/F(x)$ is increasing with increasing x. Log $F(x)$ is therefore a convex function, and therefore $F(x)$, for the convexity of a function follows from the convexity of its logarithm as can be easily verified.

Since $F'(x)<0$, $F(x)$ is a decreasing function of x, and we see by (1) that

[1] Presented to the Society, April 13, 1940.

[2] A. E. Mayer, Acta Mathematica, vol. 70 (1938), pp. 57–62.

[3] Mayer, loc. cit.

$$x^{-p} > F^2(x+a) > (x+a)^{-p}. \tag{4}$$

We can now prove the main part of our theorem: if $f(x)$ is a function convex for $x > K \geqq 0$, and if it satisfies (1) for all $x > 0$, then $f(x) \equiv F(x)$. From (1) it follows that the quotient $Q(x) = f(x)/F(x)$ satisfies the functional equations

$$Q[x + (2n-1)a] = 1/Q(x), \tag{5}$$

$$Q[x + 2na] = Q(x), \tag{6}$$

where n is any positive integer. We must show that $Q(x) \equiv 1$. First we suppose that $Q(x) > 0$ for some particular value of x. Then for that value of x by (4), (5), and (6)

$$f[x + (2k-1)a] = \frac{F[x + (2k-1)a]}{Q(x)} < \frac{1}{Q(x)[x + (2k-2)a]^{p/2}}$$

$$f[x + 2ka] = Q(x)F(x + 2ka) > \frac{Q(x)}{[x + 2ka]^{p/2}}.$$

Making use of the convexity of $f(x)$ for $x > K$, we may write for $x + (2n-1)a > K$,

$$f(x + 2na) \leqq \tfrac{1}{2}\{f[x + (2n-1)a] + f[x + (2n+1)a]\}, \tag{7}$$

and therefore

$$\frac{Q(x)}{(x + 2na)^{p/2}} < \frac{1}{2Q(x)}\left\{\frac{1}{[x + (2n-2)a]^{p/2}} + \frac{1}{(x + 2na)^{p/2}}\right\},$$

or

$$Q^2(x) < \frac{1}{2}\left\{\left[\frac{x + 2na}{x + (2n-2)a}\right]^{p/2} + 1\right\}.$$

If $Q(x) > 0$, then $Q(x+a) = 1/Q(x) > 0$, and in an exactly similar way we find that

$$Q^2(x + a) = \frac{1}{Q^2(x)} < \frac{1}{2}\left\{\left[\frac{x + (2n+1)a}{x + (2n-1)a}\right]^{p/2} + 1\right\}.$$

Since the last two displayed equations are true for all values of n such that $x + (2n-1)a > K$, $Q^2(x) = 1$, and therefore $|f(x)| = F(x)$. If we assume that there exists an x for which $Q(x) < 0$, an argument similar to the preceding one shows again that $|f(x)| = F(x)$. Hence we conclude that for all values of $x > 0$ $|f(x)| \equiv F(x)$. We thus see, that, for every value of x, $f(x)$ is equal to either plus or minus $F(x)$. If $f(x)$ were

equal to minus $F(x)$ for some particular value of x, say $x=x_0$, then by (1) $f(x_0+na)<0$ for all n. Hence there would exist an x, say x_1, larger than K such that $f(x_1)<0$, and $f(x_1)=-F(x_1)$. Then for every $h>0$

$$f(x_1+h) \leqq \tfrac{1}{2}[f(x_1)+f(x_1+2h)] \leqq \tfrac{1}{2}[-F(x_1)+F(x_1+2h)].$$

But $F(x)$ is decreasing, and therefore $f(x_1+h)$ would be negative for all $h>0$, and equal to $-F(x_1+h)$. This function is, however, not convex. Hence $f(x)$ can never be negative, and must be identical with $F(x)$.

A function is said to be concave for $x>K$ if it satisfies equation (3) with the inequality sign reversed. It is obvious that the only function which is concave for $x>K\geqq 0$ and satisfies the functional equation (1) is $-F(x)$.

If the condition of convexity be replaced by requirement that $f(x)$ be monotone for $x>K\geqq 0$, we can obtain the following result:

THEOREM 2. *The only function $f(x)$ which is monotone, not increasing, for all $x>K\geqq 0$, and satisfies the functional equation* (1) *is $F(x)$.*

The proof of this theorem is analogous to that of Theorem 1, with the convexity condition (7) replaced by the monotone relation

$$(8) \qquad f(x+2na) \leqq f[x+(2n-1)a] \leqq f[x+(2n-2)a]$$

(n sufficiently large). If $x>0$ is such that $Q(x)>0$, then from (8) it follows that $Q^2(x)=1$. And if it is supposed that $Q(x)<0$ for some x, (8) leads to the contradictory inequalities $Q^2(x)>1$, $Q^2(x)<1$. Hence $f(x)\equiv F(x)$. The details of this proof are left to the reader.

It is also obvious from the last theorem that the only function $f(x)$ which is monotone, not decreasing, for all $x>K\geqq 0$, and satisfies (1) is $-F(x)$.

COLLEGE OF ST. THOMAS

AN UNSYMMETRIC FUBINI THEOREM[1]

R. H. CAMERON AND W. T. MARTIN

Because of the great variety of ways in which repeated Stieltjes integrals may occur, the ordinary Fubini theorem for Lebesgue integrals has more than one analogue in Lebesgue-Stieltjes (Radon) integrals.[2] One of these (the symmetric one) is well known.[3] However, there is another form which is not symmetric, and its proof does not seem to be in the literature, although it is known to many writers, in special cases at least.[4] This form does not appear to be immediately derivable from the symmetric form, and since it is of interest in various connections, it seems worthwhile to give an explicit statement and proof. We do not follow the usual procedure of beginning with finite limits and then allowing the limits to become infinite because the passage to infinity seems to present difficulties of the same order as the direct proof of the final result itself. The immediate use of infinite limits is made possible by the fact that the symmetric Fubini theorem has already been proved with infinite limits.

THEOREM. *Let $k(x)$ be a function of bounded variation on every finite interval. Let $p(x, u)$ be Borel measurable in (x, u); for almost all x with respect to $k(x)$ let it be of bounded variation in u over every finite u-interval. Denote by $V(x, u)$ the variation*[5] $V(x, u) = \int_{0+0}^{u+0} |d_v p(x, v)|$. *Assume that $\int_{-\infty}^{\infty} V(x, u) |dk(x)|$ exists (is finite) for all*[6] *u. Let $s(u)$ be Borel measurable on $(-\infty, \infty)$. Then the existence (finiteness) of either*

$$\int_{-\infty}^{\infty} |s(u)| \, d_u \int_{-\infty}^{\infty} V(x, u) \, |dk(x)| \tag{1}$$

[1] Presented to the Society, February 24, 1940.

[2] The integrals which occur in this paper are understood to be Lebesgue-Stieltjes (Radon) integrals. See, for example, Saks, *Theory of the Integral*, 2d revised edition, Warsaw-Lemberg, 1937, pp. 19 and 67. We note that the familiar principles of monotone and dominated convergence in the Lebesgue theory are also valid in this theory; see pages 28 and 29. In this paper we shall not admit $\pm \infty$ as members of our number system; that is, existence implies finiteness.

[3] Saks, loc. cit., p. 81.

[4] See for instance, N. Wiener and H. R. Pitt, *On absolutely convergent Fourier-Stieltjes transforms*, Duke Mathematical Journal, vol. 4 (1938), pp. 420–436. The reader will note that a form of this theorem was used in passing from line 13 to line 14 on page 421.

[5] For negative u we understand that $V(x, u) = -\int_{u+0}^{0+0} |d_v p(x, v)|$. Of course V is strictly a variation only for positive u.

[6] Clearly it would be sufficient to require that the integral be finite for u ranging over some sequence having $\pm \infty$ as limit points.

or

$$\int_{-\infty}^{\infty} |\, dk(x)\,| \int_{-\infty}^{\infty} |\, s(u)\,| \cdot |\, d_u p(x, u)\,| \tag{2}$$

implies the existence of the other, their equality, and the existence and equality of the following integrals

$$\int_{-\infty}^{\infty} s(u) d_u \int_{-\infty}^{\infty} p(x, u) dk(x) = \int_{-\infty}^{\infty} dk(x) \int_{-\infty}^{\infty} s(u) d_u p(x, u). \tag{3}$$

We shall first prove this theorem under the following additional assumptions which we shall later successively remove:

(A) $s(u)$ is a polygonal function (that is, a continuous function whose second derivative exists and is zero except at a finite number of points),

(B) $s(u)$ is bounded and vanishes identically outside a finite interval,

(C) $s(u)$ is non-negative, $k(x)$ is monotonic increasing and $p(x, u)$ is monotonic increasing and right continuous in u and $p(x, 0) = 0$.

Clearly the additional hypotheses (B) and (C) imply the existence of (1). Also they imply the existence of (2), as one sees in the following manner. Let n be such that $s(u) = 0$ when $|u| > n-1$ and such that $s(u)$ is everywhere less than or equal to n. Then since $V(x, u) = p(x, u)$ by (C) we see that

$$\int_{-\infty}^{\infty} s(u) d_u p(x, u) \leqq n \int_{-n}^{n} d_u p(x, u) = n[V(x, n) - V(x, -n)]. \tag{4}$$

Also

$$\int_{-\infty}^{\infty} V(x, u) dk(x) \tag{5}$$

exists by hypothesis. Hence

$$\int_{-\infty}^{\infty} dk(x) \int_{-\infty}^{\infty} s(u) d_u p(x, u) \leqq n \int_{-\infty}^{\infty} dk(x)[V(x, n) - V(x, -n)];$$

that is, (2) also exists.

Thus under the additional hypotheses (B) and (C) we need only show the equality of (1) and (2) since obviously the right and left members of (3) are equal to (1) and (2) respectively. In order to show the equality of (1) and (2) under (A), (B) and (C) we integrate by parts and use the fact that s vanishes at $-n$ and at n. Thus

$$(6) \qquad \int_{-n}^{n} s(u)d_u \int_{-\infty}^{\infty} p(x, u)dk(x) = -\int_{-n}^{n} ds(u) \int_{-\infty}^{\infty} p(x, u)dk(x).$$

By the symmetric Fubini theorem, we obtain

$$(7) \qquad \int_{-n}^{n} ds(u) \int_{-\infty}^{\infty} p(x, u)dk(x) = \int_{-\infty}^{\infty} dk(x) \int_{-n}^{n} p(x, u)ds(u).$$

Integrating the inner integral of the right member of (7) by parts, we have

$$(8) \qquad \int_{-\infty}^{\infty} dk(x) \int_{-n}^{n} p(x, u)ds(u) = -\int_{-\infty}^{\infty} dk(x) \int_{-n}^{n} s(u)d_u p(x, u)$$

and this with (6) and (7) establishes the equality of (1) and (2) and thus in view of the preceding remarks proves the theorem under the additional assumptions (A), (B) and (C).

Now since every bounded Borel measurable function is a repeated limit of continuous (and even of polygonal) functions[7] we shall show by the repeated use of the principle of dominated convergence that the theorem still holds when we drop (A) and keep only the additional hypotheses (B) and (C). As we have noted, for this purpose it is again only necessary to prove the equality of (1) and (2). For purposes of induction let us assume that $s^*(u) = \lim_{j\to\infty} s_j(u)$ $(-\infty < u < \infty)$ where $s_1(u)$, $s_2(u)$, $\cdots$ is a sequence of non-negative functions each of which is Borel measurable and vanishes when $|u| > n-1$ and each of which is everywhere less than or equal to n. Assuming that the conclusion of the theorem holds for each $s_j(u)$ when p and k satisfy (C), we shall show that it also holds for the limit function $s^*(u)$. By bounded convergence

$$\lim_{j\to\infty} \int_{-n}^{n} s_j(u)d_u \int_{-\infty}^{\infty} p(x, u)dk(x) = \int_{-n}^{n} s^*(u)d_u \int_{-\infty}^{\infty} p(x, u)dk(x);$$

and since (4) holds for each $s_j(u)$ it follows by dominated convergence that

$$\lim_{j\to\infty} \int_{-\infty}^{\infty} dk(x) \int_{-n}^{n} s_j(u)d_u p(x, u) = \int_{-\infty}^{\infty} dk(x) \lim_{j\to\infty} \int_{-n}^{n} s_j(u)d_u p(x, u)$$

$$= \int_{-\infty}^{\infty} dk(x) \int_{-n}^{n} s^*(u)d_u p(x, u).$$

[7] See, for example, de la Vallée Poussin, *Intégrales de Lebesgue* (Borel Monograph), 1916, pp. 36, 37.

Thus if the theorem holds under (B) and (C) for one Baire class it holds under (B) and (C) for the next; and hence the theorem holds under (B) and (C) for all Borel measurable functions.

In order to remove (B) (retaining only (C)) let

$$s^n(u) = \begin{cases} s(u) & \text{if } |u| \leqq n-1 \text{ and } s(u) \leqq n, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Then since $s^n(u)$ satisfies (B) and (C),

$$\int_{-\infty}^{\infty} s^n(u) d_u \int_{-\infty}^{\infty} p(x, u) dk(x) = \int_{-\infty}^{\infty} dk(x) \int_{-\infty}^{\infty} s^n(u) d_u p(x, u). \tag{10}$$

Since either (1) or (2) exists, both these expressions are bounded (in n). By monotone convergence

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} s^n(u) d_u \int_{-\infty}^{\infty} p(x, u) dk(x) = \int_{-\infty}^{\infty} s(u) d_u \int_{-\infty}^{\infty} p(x, u) dk(x),$$

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} dk(x) \int_{-\infty}^{\infty} s^n(u) d_u p(x, u) = \int_{-\infty}^{\infty} dk(x) \lim_{n \to \infty} \int_{-\infty}^{\infty} s^n(u) d_u p(x, u)$$

$$= \int_{-\infty}^{\infty} dk(x) \int_{-\infty}^{\infty} s(u) d_u p(x, u).$$

Thus the theorem holds under (C).

Finally, we can dispense with (C) and assume merely the original hypotheses of the theorem. To show this let

$$s_i(u) = \max \{(-1)^i s(u), 0\}, \qquad i = 0, 1,$$

$$k_i(x) = \frac{1}{2}\left\{ \int_0^{x+0} |dk(\xi)| + (-1)^i k(x+0) \right\}, \qquad i = 0, 1, \tag{11}$$

$$p_i(x, u) = \tfrac{1}{2}\{V(x, u) + (-1)^i [p(x, u+0) - p(x, 0+0)]\}, \qquad i = 0, 1.$$

Obviously all these functions are Borel measurable and satisfy (C). We first show that the existence of (1) or (2) for s, p and k implies the existence of (1) or (2) respectively for s_i, p_j, k_l. Since we always have

$$\int_{-\infty}^{\infty} |\phi(x)| df(x) \geqq \int_{-\infty}^{\infty} |\phi(x)| dg(x)$$

when $f(x) - g(x)$ is non-decreasing, and

$$\int_{-\infty}^{\infty} \phi(x) \, | \, df(x) \, | \geqq \int_{-\infty}^{\infty} \theta(x) \, | \, df(x) \, |$$

when $\phi(x) \geqq \theta(x)$, it follows that

$$\begin{aligned}\int_{-\infty}^{\infty} | \, s(u) \, | \, d_u \int_{-\infty}^{\infty} V(x, u) \, | \, dk(x) \, | &\geqq \int_{-\infty}^{\infty} s_i(u) d_u \int_{-\infty}^{\infty} V(x, u) \, | \, dk(x) \, | \\ &\geqq \int_{-\infty}^{\infty} s_i(u) d_u \int_{-\infty}^{\infty} p_j(x, u) \, | \, dk(x) \, | \\ &\geqq \int_{-\infty}^{\infty} s_i(u) d_u \int_{-\infty}^{\infty} p_j(x, u) dk_l(x)\end{aligned}$$

and

$$\begin{aligned}\int_{-\infty}^{\infty} | \, dk(x) \, | \int_{-\infty}^{\infty} | \, s(u) \, | \, | \, d_u p(x, u) \, | &\geqq \int_{-\infty}^{\infty} dk_l(x) \int_{-\infty}^{\infty} | \, s(u) \, | \, | \, d_u p(x, u) \, | \\ &\geqq \int_{-\infty}^{\infty} dk_l(x) \int_{-\infty}^{\infty} s_i(u) \, | \, d_u p(x, u) \, | \\ &\geqq \int_{-\infty}^{\infty} dk_l(x) \int_{-\infty}^{\infty} s_i(u) d_u p_j(x, u).\end{aligned}$$

Thus s_i, p_j, k_l satisfy the hypotheses of the theorem, as well as (C), and hence

$$\begin{aligned}(-1)^{i+j+l} \int_{-\infty}^{\infty} s_i(u) d_u \int_{-\infty}^{\infty} p_j(x, u) dk_l(x) \\ = (-1)^{i+j+l} \int_{-\infty}^{\infty} dk_l(x) \int_{-\infty}^{\infty} s_i(u) d_u p_j(x, u).\end{aligned}$$

Adding these eight equations yields (3), and hence completes the proof of the theorem.

THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINEAR FORMS IN FUNCTION FIELDS[1]

LEONARD TORNHEIM

We shall prove algebraically an analogue for function fields[2] of a well known theorem of Minkowski on linear forms.[3]

THEOREM 1. *Let F be a field and z an indeterminate over F. Let*

$$L_i = \sum_{j=1}^{n} a_{ij}x_j, \qquad i = 1, \cdots, n, \tag{1}$$

be n linear expressions with coefficients a_{ij} in $F(z)$ and with the determinant $|a_{ij}|$ of degree[4] d. Then for any set of n integers $c_1, \cdots, c_n$ which satisfy the condition $\sum_{i=1}^{n} c_i > d-n$ there exists a set of values for $x_1, \cdots, x_n$ in $F[z]$ and not all zero such that each L_i has degree at most c_i.

First, we may assume that all of the c_i are equal. For, suppose that c is the maximum of the c_i. Write L_i' for $L_i z^{c-c_i}$. The determinant of the coefficients of the L_i' has degree $d' = d + \sum(c-c_i) < \sum c + n$. If there is a set of values for $x_1, \cdots, x_n$ with the property that the degree of each L_i' is at most c, then these same values will make the degree of L_i at most c_i.

Next, we may assume, after multiplying each L_i by a suitable polynomial and by using an argument similar to that above, that all the a_{ij} are in $F[z]$.

We shall now convert our system of L_i by means of a transformation of determinant unity with elements in $F[z]$ into an equivalent system having $a_{ij}=0$ for $i<j$. Let b_1 be the g.c.d. of the a_{1j}; then $b_1 = \sum_{j=1}^{n} a_{1j}c_{j1}$ for appropriate c_{j1} in $F[z]$. Necessarily the c_{j1} are relatively prime. It is possible to find other quantities c_{jk} $(k=2, \cdots, n)$ such that the determinant $|c_{jk}|$ has value unity.[5] Thus the transfor-

[1] Presented to the Society, April 13, 1940.

[2] See M. Deuring, *Zur Theorie der Idealklassen in algebraischen Funktionenkörpern*, Mathematische Annalen, vol. 106 (1932), pp. 103–106, for a related result. I believe the results I prove are new.

[3] A bibliography of both analytic and algebraic proofs of the theorem of Minkowski on linear forms is given by E. Jacobsthal, *Der Minkowskische Linearformensatz*, Sitzungsberichte Berliner mathematischen Gesellschaft, vol. 33 (1934), pp. 62–64. See also L. J. Mordell, *Minkowski's theorem on homogeneous linear forms*, Journal of the London Mathematical Society, vol. 8 (1933), pp. 179–192.

[4] The degree of a rational function is the degree of the numerator less that of the denominator. Zero is assigned the degree minus infinity.

[5] A. A. Albert, *Normalized integral bases of algebraic number fields* I, Annals of Mathematics, (2), vol. 38 (1937), p. 926 ff. The statement is proved for rational integral c_{jk} but the proof applies to any integral domain having the property that a

mation $x_j=\sum_{k=1}^{n}c_{jk}x_k'$ has determinant unity and hence it has a reciprocal transformation with elements in $F[z]$. The forms L_i are transformed into $L_i'=\sum_{k=1}^{n}a_{ik}'x_k'$. Here $a_{ik}'=\sum_{j=1}^{n}a_{ij}c_{jk}$, and, being a linear combination of a_{1j}, it is divisible by their g.c.d. b_1; $a_{ik}'=b_1a_k$. The transformation

$$x_1' = x_1' - \sum_{k=2}^{n} a_k x_k'', \qquad x_r' = x_r'', \qquad r = 2, \cdots, n,$$

of determinant unity transforms the L_i' into L_i'' with $L_1''=b_1x_1''$.

The procedure is repeated for the $n-1$ linear forms $M_i=\sum_{j=2}^{n}a''_jx_j''$ $(i=2, \cdots, n)$. Finally, if this process is continued, the resultant transformation converts the original system (1) into one with $a_{ij}=0$ for $i<j$. As a consequence, if the degree of a_{ii} is d_i, then $\sum d_i=d$. By using another transformation of determinant unity we may assume that the degree of each a_{ij} is at most d_i.

Let G_1 be the set of all n-tuples $(s_1, \cdots, s_n)=s$ where the s_i are in $F[z]$ and have degree not greater than c; hence G_1 is a linear set over F whose order $u_1=n(c+1)$. Write $L_i(s)$ for $\sum_{j=1}^{n}a_{ij}s_j$. Let G_r of order u_r over F be the linear subset of G_1 composed of all quantities s for which $L_1, \cdots, L_{r-1}$ all take values of degree not greater than c. Designate by P_r the set of all $L_r(s)$ with s in G_r, and by Q_r the set of all polynomials in P_r of degree not exceeding c. Since the maximum degree possible for a polynomial in P_r is $c+d_r$, the number of linearly independent polynomials of P_r which are not in Q_r, that is, the order of P_r/Q_r, is less than or equal to d_r. Now $G_r/G_{r+1}\simeq P_r/Q_r$, a fact which follows from the mapping of G_r on P_r and G_{r+1} on Q_r. Hence $[G_1:G_{n+1}]\leqq\sum_{i=1}^{n}d_i=d$. Therefore the order u_{n+1} of G_{n+1} is not less than $n(c+1)-d$. To be sure that G_{n+1} has elements other than zero, we must have $u_{n+1}\geqq 1$, that is, $nc=\sum c\geqq d+1-n$.

The following theorem applies if some of the L_i must be made equal to zero.

THEOREM 2. *If in Theorem* 1 *the first m of the L_i are to be made equal to zero and if their coefficients are in $F[z]$, then the conclusion will hold if $\sum_{i=m+1}^{n}c_i>d-(n-m)$.*

For, the first m polynomials s_i must be zero if we have the transformed system used in the proof of Theorem 1. Application of Theorem 1 for the remaining L_i yields Theorem 2.

UNIVERSITY OF CHICAGO

g.c.d. of any finite number of elements exists and is linearly expressible in terms of those elements, that is, that every ideal with a finite basis is principal.

A NOTE ON A THEOREM OF RADÓ CONCERNING THE $(1, m)$ CONFORMAL MAPS OF A MULTIPLY-CONNECTED REGION INTO ITSELF[1]

MAURICE H. HEINS

Let G_z denote a region in the z-plane and let $w=f(z)$ be a function of z defined for $z \subset G_z$ which has the following properties: (1) $w=f(z)$ is analytic and single-valued for $z \subset G_z$, (2) $z \subset G_z$ implies that $f(z) \subset G_z$, (3) to each point $w_0 \subset G_z$ there correspond m and only m points $z_0^{(k)}$ $(k=1, 2, \cdots, m)$ contained in G_z such that $f(z_0^{(k)}) = w_0$ $(k=1, 2, \cdots, m)$ where following the usual convention we count the $z_0^{(k)}$ according to their multiplicities. Then $w=f(z)$ is said to define a $(1, m)$ *conformal map of* G_z *onto itself*. Such maps have been studied by Fatou[2] and Julia[3] for the case where G_z is simply-connected, and by Radó[4] who treated multiply-connected regions as well. Among the results which Radó established is the following theorem:

Let G_z be a region of finite connectivity p (>1); then there exists no $(1, m)$ conformal map of G_z onto itself for $m>1$.

Let us remark with Radó that the theorem is no longer valid if G_z is of infinite connectivity, as simple examples from the theory of the iteration of rational functions show.[5] Radó's proof of the theorem just cited is based on the possibility of mapping one-to-one and conformally a region of finite connectivity p, none of the components of its boundary reducing to points, onto a region of connectivity p, the boundary of which consists of p disjoint circles. Other types of canonical regions yield the same result, notably one due to Koebe.[6] It is the object of the present note to establish Radó's theorem directly without appeal to the possibility of mapping one-to-one and conformally the region G_z onto a canonical region. Our tools are the theory of iteration and a simple modification of Nevanlinna's principle of harmonic measure.[7]

Let G_z, the region we are going to study, have as its boundary p (>1) disjoint continua Γ_k $(k=1, 2, \cdots, p)$. It is evident that we

[1] Presented to the Society, April 27, 1940, under the title *A note on a theorem of Radó*.

[2] P. Fatou, Bulletin de la Société Mathématique de France, 1919.

[3] G. Julia, Comptes Rendus de l'Académie des Sciences, Paris, vol. 166 (1918).

[4] T. Radó, Acta Szeged, vol. 1 (1922).

[5] G. Julia, Journal de Mathématiques Pures et Appliquées, 1918.

[6] P. Koebe, Acta Mathematica, vol. 41.

[7] R. Nevanlinna, *Eindeutige analytische Funktionen*, chap. 3.

may assume without loss of generality that the Γ_k are all closed Jordan curves. Further let $w=f(z)$ define a $(1, m)$ conformal map of G_z onto itself. In Radó's paper cited above, it is shown that if $w=f(z)$ defines a $(1, m)$ conformal map of G_z onto itself, then whenever z tends to the boundary of G_z, so does $f(z)$, in such a manner that whenever z tends to a given component Γ_k of the boundary, then $f(z)$ tends to one and only one component of the boundary Γ_{l_k} where the index l_k depends on k. Now the relation $k \to l_k$ $(k=1, 2, \cdots, p)$ defines a permutation of the indices $1, 2, \cdots, p$.

We shall understand by $f_n(z)$, *the nth iterate of* $f(z)$, that function defined by the recursive relations

$$\text{(A)} \qquad f_0(z) \equiv z,\ f_1(z) \equiv f(z), \cdots,\ f_n(z) \equiv f[f_{n-1}(z)].$$

By the definition of $w=f(z)$ it is clear that the definition given by (A) for $f_n(z)$ is meaningful. Since the relation $k \to l_k$ defines a permutation of the indices $1, 2, \cdots, p$, a suitably chosen power of this permutation is the identity. Hence for a properly chosen whole number n_0, $f_{n_0}(z)$ has the property that when z tends to a component Γ_k of the boundary of G_z, then $f_{n_0}(z)$ tends to exactly the same component Γ_k. Let us denote $f_{n_0}(z)$ by $F(z)$.

We are now in a position to demonstrate Radó's theorem directly. By the *harmonic measure of z with respect to* Γ_k, denoted by $\omega(z, \Gamma_k, G_z)$ we understand that harmonic function defined for $z \subset G_z$ which is single-valued and bounded for $z \subset G_z$ and further takes on the boundary value 1 on Γ_k and the boundary value 0 on all the components of the boundary of G_z exclusive of Γ_k.[8]

The relations

$$\text{(B)} \qquad \omega(F(z), \Gamma_k, G_z) \equiv \omega(z, \Gamma_k, G_z), \qquad k = 1, 2, \cdots, p,$$

are an immediate consequence of the principle of the maximum for harmonic functions. For, as z tends to a given component of the boundary of G_z, $F(z)$ tends to the same component. Hence $\omega(F(z), \Gamma_k, G_z)$ has the same boundary values as $\omega(z, \Gamma_k, G_z)$, hence the identity. To establish Radó's theorem we need consider only one of these identities, say

$$\text{(C)} \qquad \omega(F(z), \Gamma_1, G_z) \equiv \omega(z, \Gamma_1, G_z).$$

Let z_0 be a point lying in G_z, and let $\omega(z_0, \Gamma_1, G_z)=\lambda_0$ $(0<\lambda_0<1)$. Then our identity (C) implies

$$\omega(F_n(z_0), \Gamma_1, G_z) = \lambda_0,$$

[8] R. Nevanlinna, ibid.

where $F_n(z)$ is the nth iterate of $F(z)$ for $n=1, 2, \cdots$. Hence if z is on a given level curve defined by $\omega(z, \Gamma_1, G_z)=\lambda$ $(0<\lambda<1)$, $F_n(z)$ is on the same level curve for $n=1, 2, \cdots$. This permits us to conclude that no limit function of the sequence $\{F_n(z)\}$ is a constant. From this fact we infer that $w=F(z)$ defines a (1, 1) conformal map of G_z onto itself, and hence so does $w=f(z)$; that is, *m cannot exceed one.*

Suppose contrary to our assertion that $m>1$. Then $w=F(z)$ would define a $(1, m^{n_0})$ conformal map of G_z onto itself and $w=F_n(z)$ would define a $(1, m^{n n_0})$ conformal map of G_z onto itself. Let $\{F_{k_n}(z)\}$ be a convergent subsequence of $\{F_n(z)\}$ and let $F_0(z)$ denote the limit function of this subsequence. A point w_0 has, under the map defined by $w=F_{k_n}(z)$, $m^{k_n n_0}$ antecedents all of which lie on the level curve defined by $\omega(z, \Gamma_1, G_z)=\omega(w_0, \Gamma_1, G_z)$. But by Hurwitz's theorem, for k_n sufficiently large $F_{k_n}(z)-w_0$ and $F_0(z)-w_0$ have the same number of zeros, and this is manifestly impossible. Hence m cannot exceed one.

It is interesting to note that the technique employed in the present proof of Radó's theorem, that is, a combined use of the theory of iteration and of an extended form of the principle of harmonic measure (the modification consisting in the fact that we did not require the continuity of $F(z)$ on Γ_1) may be applied to other problems in the theory of the conformal mapping of multiply-connected regions. In particular, such a technique permits us to conclude that the number of (1, 1) conformal maps of a given multiply-connected region of finite connectivity p, where $p>2$, bounded by p disjoint continua is finite. Koebe's proof of this theorem is based on the fact that the region in question can be mapped one-to-one and conformally onto a canonical region. We shall not give in the present note the details of a direct proof of this theorem. Let us remark however that it is based on the fact that for $p>2$ each of the harmonic functions $\omega(z, \Gamma_k, G_z)$ $(k=1, 2, \cdots, p)$ has precisely $p-2$ (>0) critical points. A study of the level curves containing these critical points yields the desired result.

HARVARD UNIVERSITY

NON-CYCLIC ALGEBRAS OF DEGREE FOUR AND EXPONENT TWO WITH PURE MAXIMAL SUBFIELDS[1]

ROY DUBISCH

In a recent paper[2] A. A. Albert proved the falsity of the converse of the well known proposition that a cyclic normal division algebra contains a quantity j whose minimum equation is $x^n = j$ in the base field of the algebra. The proof consists in giving an example of a non-cyclic normal division algebra containing a quantity j as described above. The algebra described in that example was of degree and exponent four. It is the purpose of this paper to show that the exponent does not affect the property, and this we shall do by constructing similarly an algebra of degree four whose exponent is two.

We shall actually prove the following theorem.

THEOREM. *Let ξ and η be independent indeterminants over the field R of real numbers, $K = R(\xi, \eta)$. Then there exist non-cyclic normal division algebras of degree four and exponent two over K such that $t^4 = \gamma$ in K, t^2 not in K, for some quantity t in each algebra A.*

To make our proof we use the known[3] property stating that a normal division algebra of degree four has exponent two if and only if it is expressible as a direct product of two algebras of degree two. Therefore we may take our desired algebra A to be

$$A = B \times C = (1, i, j, ij) \times (1, x, y, xy),$$

$$(1) \qquad ji = -ij, \quad i^2 = u, \quad j^2 = a, \qquad u \neq 0,\ a \neq 0 \text{ in } K,$$

$$yx = -xy, \quad x^2 = v, \quad y^2 = b, \qquad v \neq 0,\ b \neq 0 \text{ in } K.$$

We seek first a quantity t with minimum equation $t^4 = \gamma$ in K. Now if we take $t = a_1 i + a_2 j + a_3 ij$, where a_1, a_2, a_3 are in $K(x)$, and if we put

$$(2) \qquad a_1 = c_1 + c_2 x, \quad a_2 = d_1 + d_2 x, \quad a_3 = f_1 + f_2 x, \quad c_i, d_i, f_i \text{ in } K,$$

we can easily compute that if

$$(3) \qquad a = -\frac{c_1^2 u + c_2^2 uv}{d_1^2 + d_2^2 v - f_1^2 u - f_2^2 uv},$$

[1] Presented to the Society, April 13, 1940.

[2] A. A. Albert, *Non-cyclic algebras with pure maximal subfields*, this Bulletin, vol. 44 (1938), pp. 576–579.

[3] A. A. Albert, *Normal division algebras of degree four*, Transactions of this Society, vol. 34 (1932), p. 369.

and if

$$b_2 = 2c_1c_2u + 2d_1d_2a - 2f_1f_2ua \neq 0, \tag{4}$$

the algebra A defined by (1) contains a quantity t such that $t^4=\gamma$ in K, t^2 not in K.

In proving A a non-cyclic division algebra we use the following method.[4] Let $z^2=\delta^2+\epsilon^2=\Delta$ in K, δ and ϵ in K with $L=K(z)$ a quadratic field over K. The algebras $B_0=B\times L$ and $C_0=C\times L$ over L are generalized quaternion algebras over their reference field L. Furthermore, $A_0=A\times L$ is the direct product $A_0=B_0\times C_0$. Then it is known[5] that a necessary and sufficient condition that A_0 over L shall be a division algebra is that the quadratic form

$$Q \equiv u\lambda_1^2 + a\lambda_2^2 - ua\lambda_3^2 - (v\lambda_4^2 + b\lambda_5^2 - vb\lambda_6^2) \tag{5}$$

in the variables $\lambda_1, \lambda_2, \cdots, \lambda_6$ in L shall not vanish for any $\lambda_1, \cdots, \lambda_6$ not all zero in L.

If $J=R[\xi, \eta]$ is the integral domain of all polynomials in ξ and η with real coefficients, it is obvious that we may, without loss of generality, take the λ_i in (5) to be in J. Hence we may write

$$\lambda_i = \alpha_i + \beta_i z, \qquad i = 1, \cdots, 6, \tag{6}$$

where the α_i and β_i are in J. Then $\lambda_i^2=(\alpha_i^2+\beta_i^2\Delta)+2\alpha_i\beta_i z$ so that if

$$P_i = \alpha_i^2 + \beta_i^2, \qquad Q_i = 2\alpha_i\beta_i, \tag{7}$$

the equation $Q\equiv 0$ becomes

$$\begin{aligned} &uP_1 + aP_2 - uaP_3 - vP_4 - bP_5 + vbP_6 \\ &\quad + (uQ_1 + aQ_2 - uaQ_3 - vQ_4 - bQ_5 + vbQ_6)z \equiv 0. \end{aligned} \tag{8}$$

But 1 and z are linearly independent[6] with respect to K so that (8) implies that

$$\phi(\xi, \eta) \equiv uP_1 + aP_2 - uaP_3 - vP_4 - bP_5 - vbP_6 \equiv 0. \tag{9}$$

Now if in (3) we let $c_1=c_2=d_1=1$, $d_2=f_1=f_2=0$, we obtain

$$a = -(u + uv). \tag{10}$$

---

[4] This is the device used by A. A. Albert in his paper *A construction of non-cyclic normal division algebras*, this Bulletin, vol. 38 (1932), pp. 452–453.

[5] A. A. Albert, *On the Wedderburn norm condition for cyclic algebras*, this Bulletin, vol. 37 (1931), p. 311, Theorem 3.

[6] It is obvious that we can have this condition hold by a suitable choice of δ and ϵ.

Note that these choices also satisfy (4). Substituting this value for a in (3) we have

$$(11)\quad \phi(\xi, \eta) = u(P_1 - P_2) - uvP_2 + u^2P_3 + u^2vP_3 - vP_4 - bP_5 + vbP_6.$$

With these choices for the parameters of our algebra A is easy to prove by the well known method of an argument[7] on the degrees of the polynomial in (7) that $\phi(\xi, \eta)$ cannot be identically zero unless all of the coefficients are zero. From this it follows that A_0 over L is a division algebra over L; so is algebra A a division algebra over K. By its form it is a normal division algebra of degree four and exponent two over K.

The proof of the non-cyclic character of A is exactly the same as that given by A. A. Albert in a previous paper,[8] and with this we have the desired theorem.

WILSON JUNIOR COLLEGE

[7] A similar argument was employed by A. A. Albert in the paper last cited.

[8] This Bulletin, vol. 38 (1932), p. 454.

NOTE ON NORMALITY IN QUASI-GROUPS

D. C. MURDOCH

An interesting problem in the theory of quasi-groups is to determine how strong an associative law must be assumed in order to obtain a theory of normal subquasi-groups similar to that of ordinary groups. The properties of normal subgroups which it is desirable to retain in the non-associative case are, first, that they form a Dedekind structure, and second, that each one gives rise to a quotient group homomorphic to the whole group. In this note we shall take the latter property as the definition of normality and show that the former follows from it under very general conditions.

We shall understand by a quasi-group G a system of elements satisfying the following two postulates:

I. PRODUCT AXIOM. *Any ordered pair of elements a, b of G has a unique product ab which also belongs to G.*

II. QUOTIENT AXIOM. *For any two elements a and b of G there exist unique elements x and y of G such that*

$$ax = b, \qquad ya = b.$$

We shall make one further assumption, namely that there exist left coset expansions with respect to any subquasi-group H of G. In other words, if H is any subquasi-group, the cosets aH and bH are either identical or have no elements in common. This assumption has been shown by Hausmann and Ore [1] to be equivalent to the following:

III. WEAK ASSOCIATIVE LAW. *If a and b are arbitrary elements of G and c_0 a fixed element, let d_0 be determined so that $(ab)c_0 = ad_0$. Then for any c*

$$(ab)c = ad,$$

where d belongs to the quasi-group $\{c_0, d_0, c\}$ generated by c_0, d_0 and c.

Since no confusion can arise, we shall for convenience use the term subgroup for subquasi-group and quotient group for quotient quasi-group, without implying thereby that the systems in question are associative.

It should be noted that a theory of normal subgroups has been given by Hausmann and Ore [1] based on the definition of a normal subgroup as a subgroup H such that $aH = Ha$ for all a in G. This

theory requires that G should satisfy associative laws considerably stronger than III in order that the normal subgroups so defined shall have the desired properties. To overcome this difficulty we shall adopt instead the following:

DEFINITION: *A subgroup H of G shall be called normal if*

$$(aH)(bH) = (ab)H,$$

for all elements a and b of G.

This is the same definition used in a previous paper by the author [2] where it was applied, however, only to quasi-groups satisfying rather special associative laws. It will be shown here that the results of that paper concerning normal subgroups can be extended to any quasi-group satisfying III.

Comparing this definition with that of Hausmann and Ore, it is seen that neither one implies the other. Under the latter all subgroups of a commutative quasi-group are normal, while under the former this property is enjoyed by the "abelian" quasi-groups [2] which satisfy the law $(ab)(cd) = (ac)(bd)$. In a Suschkewitsch quasi-group [3] (having unique right unit e and satisfying $a(bc) = (ab)(ec)$, the right unit e is a subgroup which is normal in the sense of this paper but not in the sense of Hausmann and Ore.

Examples of quasi-groups satisfying III but not satisfying the stronger laws previously discussed can easily be constructed by rearranging the elements in the Cayley square of a group, or known quasi-group, so as to destroy the associativity without disturbing the existence of cosets. The following, for example,

	1	2	3	4	5	6	7	8	9
1	1	3	2	5	6	4	9	7	8
2	3	2	1	6	4	5	7	8	9
3	2	1	3	4	5	6	8	9	7
4	6	4	5	7	8	9	2	3	1
5	4	5	6	8	9	7	3	1	2
6	5	6	4	9	7	8	1	2	3
7	8	9	7	3	1	2	6	5	4
8	9	7	8	1	2	3	4	6	5
9	7	8	9	2	3	1	5	4	6

satisfies III since cosets exist with respect to all subgroups. It is easily verified that it satisfies none of the laws previously discussed by the author nor the laws A_2, A_3, A_4 or A_5 of Hausmann and Ore.

It is natural to require next that every element of a coset aH shall define the same coset aH, or in other words that $(ah)H = aH$ for all h in H. In this connection we have the following:

THEOREM 1. *The necessary and sufficient condition that every coset aH may be defined by any one of its elements is that H contain the right unit of every element of G.*

PROOF. If H contains all right units of G, then $aH = (ah)H$ for all h in H, since each coset contains the element ah. Conversely if $aH = (ah)H$ for all h in H and all a in G, then ah must belong to $(ah)H$. Hence H contains the right unit of ah for all a in G and therefore contains all right units of G.

The main object of this note is to prove that in any finite quasigroup the normal subgroups which contain all right units of G form a Dedekind structure. With a slight strengthening of the definition of normality the result may also be proved in the infinite case. We shall assume for the present, however, that G is finite.

THEOREM 2. *If H and K are normal subgroups of G and contain all right units of G, their crosscut $H \cap K = D$ is also normal.*

PROOF. From the normality of H and K we have

$$(aD)(bD) \subseteq (ab)D,$$

and since D is finite the equality must hold. Moreover D is not void since it contains all right units of G.

THEOREM 3. *If H and K are normal and contain all right units of G, their union $H \cup K$ consists of all products hk where h belongs to H and k to K.*

The proof of this theorem is similar to that of Theorem 7 of the author's previous paper already cited. The union $H \cup K$ may therefore be written as the product HK.

THEOREM 4. *If H and K are normal, and contain all right units of G, their union is also normal.*

PROOF. Using the notation e_a for the right unit of a, we have, from the normality of H and K,

$$\begin{aligned}[a(HK)][b(HK)] &= [(ae_a)(HK)][(be_b)(HK)] \\ &= [(aH)K][(bH)K] = [(aH)(bH)]K = [(ab)H]K \\ &= [(ab)e_{ab}][HK] = (ab)(HK),\end{aligned}$$

which completes the proof.

THEOREM 5. *The normal subgroups of G which contain all right units of G form a Dedekind structure.*

PROOF. From Theorems 2 and 4 these subgroups form a structure. It is only necessary, therefore, to show that if $M \supset H$,

$$M \cap (H \cup K) \subset H \cup (M \cap K).$$

This follows as usual since every element of the left-hand side has the form hk and belongs to M. Hence k belongs to $M \cap K$ since h belongs to M.

In establishing Theorem 5 we have used the finiteness of G in the proofs of Theorems 2, 3 and 4. This can be avoided by strengthening the definition of normality as follows:

A subgroup H shall be called normal in G if for all elements a, b of G and every element h of H we have

$$(ah)(bH) = (ab)H, \qquad (aH)(bh) = (ab)H.$$

This of course implies the other definition and is equivalent to it in case G is finite. This strengthening of the definition enables us to prove the quotient axioms for the union and crosscut in Theorems 2 and 3. For example, for the union, the equation $x(h_1k_1) = h_2k_2$ has a unique solution of the form $x = hk$. For choosing h so that $hh_1 = h_2$, it follows from (2) that there exists an element k of K such that

$$(hk)(h_1k_1) = h_2k_2.$$

For the proof of Theorem 4 we have as in the finite case

$$a(HK) = (ae_a)(HK) \subset (aH)K.$$

But in the infinite case it is necessary to show also that $(aH)K \subset a(HK)$. This follows from III, applied in the form of Theorem 7, Chapter I, Hausmann and Ore.

From Theorem 5 the usual theorems concerning normal subgroups follow. In particular, the Jordan-Hölder theorem for principal chains, and the Schmidt-Remak decomposition theorems, as proved by Ore [4] for arbitrary Dedekind structures, will all hold. One interesting question arises which has no counterpart in ordinary groups, namely whether the subgroup R, generated by the right units of G, is necessarily normal. Although there seems to be no obvious reason to believe that this is true, I have no counterexample. If R is not always normal the question arises whether infinite quasi-groups exist in which the Dedekind structure of normal subgroups has no unit element.

BIBLIOGRAPHY

1. B. A. Hausmann and O. Ore, *Theory of Quasi-groups*, American Journal of Mathematics, vol. 59 (1937), pp. 983–1004.

2. D. C. Murdoch, *Quasi-groups which satisfy certain generalized associative laws*, American Journal of Mathematics, vol. 61 (1939), pp. 509–522.

3. A. Suschkewitsch, *On a generalization of the associative law*, Transactions of this Society, vol. 31 (1929), pp. 204–214.

4. O. Ore, *On the foundation of abstract algebra* I and II, Annals of Mathematics, vol. 36 (1935), pp. 406–437, and vol. 37 (1936), pp. 265–292.

YALE UNIVERSITY

ON THE FIRST CASE OF FERMAT'S LAST THEOREM

D. H. AND EMMA LEHMER

In 1909 Wieferich [1] proved his celebrated criterion for the first case of Fermat's last theorem, namely:

The equation

$$x^p + y^p = z^p, \qquad x, y, z \text{ prime to } p, \tag{1}$$

has no solutions unless

$$2^{p-1} \equiv 1 \pmod{p^2}. \tag{2}$$

Since that time numerous other criteria of the form

$$m^{p-1} \equiv 1 \pmod{p^2} \tag{3}$$

have been proved by Mirimanoff [2] (for $m=3$), Vandiver [3] (for $m=5$), Frobenius [4], Pollaczek [5], Morishima [6], and Rosser [7] for all prime values of $m \leqq 41$.

Wieferich's criterion alone has been applied by Meissner [8] and Beeger [9] for $p<16{,}000$ and was found to be satisfied only for $p=1{,}093$ and $3{,}511$, both of which cases failed to satisfy Mirimanoff's criterion.

Until recently no effort has been made to combine these various criteria in a practical way. Mirimanoff observed, however, in 1910 that his criterion and that of Wieferich could be combined to state that equation (1) has no solutions for all primes p of the form $2^\alpha 3^\beta \pm 1$ or $|2^\alpha \pm 3^\beta|$.

In the presence of more criteria this statement can be extended thus:

We call a number an "A_n number" (after Western) if it is divisible by no prime exceeding the nth prime p_n. If the criterion (3) has been established for all $m \leqq p_n$, then equation (1) does not hold if p is the sum or difference of two A_n numbers [10]. Since all the numbers less than p_{n+1} are A_n numbers, we may state that equation (1) has no solution for any prime in a region where the A_n numbers are so dense that they do not differ by more than $2p_{n+1}-1$. This method was used in 1938 by A. E. Western [11] to show that (1) is impossible for $16{,}000<p<100{,}000$.

A more powerful method of combining the criteria was suggested recently by Rosser [12], who observes that while the congruence

$$x^{p-1} \equiv 1 \pmod{p^2} \tag{4}$$

has only $(p-1)/2$ solutions less than $p^2/2$, every A_n number is a

solution of (4) if (1) holds and if (3) has been established as far as $m = p_n$. Hence if $\phi_n(x)$ denotes the number of A_n numbers not exceeding x, then

$$\phi_n(p^2/2) \leqq (p-1)/2, \tag{5}$$

a condition which cannot hold for small p's if the A_n numbers are sufficiently dense.

Other inequalities of this kind can be derived by separating the solutions of (4) into classes. We shall do this here only in the case in which the solutions are distinguished by their parity. For every positive odd solution $\omega < p^2/3$ of (4) there exists also a solution $(p^2-\omega)/2$ (using Wieferich's criterion) which lies between $p^2/3$ and $p^2/2$, and hence differs from ω. The number $(p-1)/2$ of solutions of (4) not exceeding $p^2/2$ is then at least the number of even A_n numbers less than $p^2/3$ plus twice the number of odd A_n numbers less than $p^2/3$. This can be written as

$$\phi_n(p^2/3) + \phi_n^*(p^2/3) \leqq (p-1)/2 \tag{6}$$

where $\phi_n^*(x)$ denotes the number of odd A_n numbers less than x.

If one is to apply an inequality such as (5) or (6) outside the limit of existing tables [13] of A_n numbers, it is necessary to find lower bounds for ϕ_n and ϕ_n^*. Rosser [12] has given a lower bound for $\phi_n(10^x)$ in the form of a polynomial $f_n(x)$ of the nth degree.

By an improved method [14] which makes use of Bernoulli polynomials we have constructed polynomials $P_n(x)$ and $Q_n(x)$ of degree n giving lower and upper bounds for $\phi_n(10^x)$, and also a polynomial $P_{n-1}^*(x)$ of degree $n-1$ giving a lower bound for $\phi_n^*(10^x)$. As we shall actually need not P_{n-1}^* but the sum $P_n + P_{n-1}^*$, we tabulate the following polynomials for $n = 13$:

$P_{13}(x)$	$P_{13}(x)+P_{12}^*(x)$	$Q_{13}(x)$
$.0^9 1380608198\ x^{13}$	$.0^9 1380608198\ x^{13}$	$.0^9 1380608198\ x^{13}$
$.0^7 1272704088\ x^{12}$	$.0^7 1326732670\ x^{12}$	$.0^7 1326732670\ x^{12}$
$.0^6 5247670000\ x^{11}$	$.0^6 5691949617\ x^{11}$	$.0^6 5729012056\ x^{11}$
$.0^4 1277269076\ x^{10}$	$.0^4 1439087748\ x^{10}$	$.0^4 1469359894\ x^{10}$
$.0^3 2039391192\ x^9$	$.0^3 2383232393\ x^9$	$.0^3 2493297644\ x^9$
$.0^2 2244543403\ x^8$	$.0^2 2716892889\ x^8$	$.0^2 2952478417\ x^8$
$.01740607057\ x^7$	$.02179700460\ x^7$	$.02510548659\ x^7$
$.09545572654\ x^6$	$.1235281024\ x^6$	$.1557243217\ x^6$
$.3654889719\ x^5$	$.4882665632\ x^5$	$.7120647453\ x^5$
$.9452666752\ x^4$	$1.302427247\ x^4$	$2.432810990\ x^4$
$1.542870198\ x^3$	$2.191273599\ x^3$	$6.376807029\ x^3$
$1.361483164\ x^2$	$1.996360212\ x^2$	$13.24421612\ x^2$
$.3459458265\ x$	$.5410860433\ x$	$21.15029718\ x$
$-.2049399029$	$-.2911608671$	19.69764203

If we now replace (5) by

$$P_{13}(\log p^2/2) \leqq (p-1)/2,$$

we find, by actual substitution into P_{13}, that this inequality holds only for $p \geqq 93{,}785{,}629$, and hence (1) has no solution[1] for $p < 93{,}785{,}629$.

Since

$$Q_{13}(\log p^2/2) \leqq (p-1)/2$$

holds only for $p > 141{,}000{,}000$, it follows that the inequality (5), even if we knew the exact value of $\phi_n(x)$, could not be used beyond this limit for $n = 13$.

The inequality (6) becomes

$$P_{13}(\log p^2/3) + P_{12}^*(\log p^2/3) \leqq (p-1)/2.$$

This holds only for $p > 102{,}108{,}200$.

Hence the first case of Fermat's last theorem is now proved for $p < 102{,}108{,}200$.

Further criteria of type (3), when established, may be used to extend this limit by calculating approximating polynomials of higher degree from the ones given above by the method described in [14].

Note added March 1. Since writing the above, Dr. Rosser has kindly sent us the manuscript of his forthcoming paper [15] in which he completes Morishima's proof that the prime 43 gives also a criterion of the form (3). We have therefore calculated $P_{14}(x)$ and $P_{13}^*(x)$ from the P_{13} and P_{12}^* given above. These polynomials are as follows:

$P_{14}(x)$	$P_{13}^*(x)$	$P_{14}(x)+P_{13}^*(x)$
$.0^{11}6037145739\ x^{14}$		$.0^{11}6037145739\ x^{14}$
$.0^{9}6683705079\ x^{13}$	$.0^{10}2544306733\ x^{13}$	$.0^{9}6938135752\ x^{13}$
$.0^{7}3337944336\ x^{12}$	$.0^{8}2536688143\ x^{12}$	$.0^{7}3591613150\ x^{12}$
$.0^{6}9940243443\ x^{11}$	$.0^{6}1131551216\ x^{11}$	$.0^{5}1107179466\ x^{11}$
$.0^{4}1965570473\ x^{10}$	$.0^{5}2980593461\ x^{10}$	$.0^{4}2263629819\ x^{10}$
$.0^{3}2719317505\ x^{9}$	$.0^{4}5152041706\ x^{9}$	$.0^{3}3234521676\ x^{9}$
$.0^{2}2700968174\ x^{8}$	$.0^{3}6140455952\ x^{8}$	$.0^{2}3315013769\ x^{8}$
$.01944029859\ x^{7}$	$.0^{2}5157919367\ x^{7}$	$.02459821796\ x^{7}$
$.1009940169\ x^{6}$	$.03064345185\ x^{6}$	$.1316374687\ x^{6}$
$.3720257684\ x^{5}$	$.1271048263\ x^{5}$	$.4991305947\ x^{5}$
$.9365601794\ x^{4}$	$.3559975936\ x^{4}$	$1.292557773\ x^{4}$
$1.500217863\ x^{3}$	$.6290315435\ x^{3}$	$2.129249406\ x^{3}$
$1.299882088\ x^{2}$	$.6014604442\ x^{2}$	$1.901342532\ x^{2}$
$.3050249350\ x$	$.1730336327\ x$	$.4780585677\ x$
$-.2081092173$	$-.08700841024$	$-.2951176274$

[1] The corresponding result obtained by Rosser is about 41,000,000.

We find that (5) is satisfied only if $p > 230$ millions, while (6) gives the better inequality $p \geqq 253{,}747{,}889$.

The first case of Fermat's last theorem is therefore now established for all $p < 253{,}747{,}889$.

Bibliography

1. A. Wieferich, Journal für die reine und angewandte Mathematik, vol. 136 (1909), pp. 293–302.

2. D. Mirimanoff, Comptes Rendus de l'Académie des Sciences, Paris, vol. 150 (1910), pp. 204–206.

3. H. S. Vandiver, Journal für die reine und angewandte Mathematik, vol. 144 (1914), pp. 314–318.

4. G. Frobenius, Sitzungsberichte der Akademie der Wissenschaften, Berlin, 1914, pp. 653–681.

5. F. Pollaczek, Sitzungsberichte der Akademie der Wissenschaften, Vienna, IIa, vol. 126 (1917), pp. 25–59.

6. T. Morishima, Japanese Journal of Mathematics, vol. 8 (1931), pp. 159–173.

7. J. B. Rosser, this Bulletin, vol. 46 (1940), pp. 299–304.

8. W. Meissner, Sitzungsberichte der Akademie der Wissenschaften, Berlin, 1913, pp. 663–667.

9. N. G. W. H. Beeger, Messenger of Mathematics, vol. 51 (1922), pp. 149–150; Nieuw Archief voor Wiskunde, (2), vol. 20 (1939), pp. 51–54.

10. This is a special case of a theorem due to Landau, l'Intermédiaire des Mathematiciens, vol. 20 (1913), p. 206. See also E. Gottschalk, Mathematische Annalen, vol. 115 (1938), pp. 157–158.

11. A. E. Western, unpublished manuscript.

12. J. B. Rosser, this Bulletin, vol. 45 (1939), pp. 636–640.

13. There is only one such table in print, namely: A. J. C. Cunningham, *Quadratic and Linear Tables*, London, 1927, pp. 162–170 (for $n = 5$). Extensive manuscript tables have been prepared by Dr. A. E. Western.

14. D. H. Lehmer, *The lattice points of an n-dimensional tetrahedron*, Duke Mathematical Journal, vol. 7 (1940), pp. 341–353.

15. J. B. Rosser, *An additional criterion for the first case of Fermat's last theorem*, this Bulletin, vol. 47 (1941), pp. 109–110.

Lehigh University

PROOF OF A THEOREM OF HALL

HARRY GOHEEN

In the Journal of the London Mathematical Society for July, 1937, Mr. Philip Hall gave a proof of the theorem, "If a group G of order g has a subgroup of order m for every divisor m of g such that $(m, g/m) = 1$, then G is a soluble group." The proof is a very simple one in contrast to the rather difficult proof of the converse theorem which Hall had published in the same journal for April, 1927. It seems worthwhile to give a simpler proof of this converse.

THEOREM. *If g is the order of a soluble group G and m is a divisor of g such that m and g/m are relatively prime, then G has a subgroup of order m and furthermore all the subgroups of order m in G are conjugate under G.*

PROOF. Since the theorem is true by default for prime power groups, let us suppose that it is true for all soluble groups of orders less than g and use the method of complete induction. Then the theorem is true for an invariant subgroup G' of prime index r in G.

If r divides g/m, then G' has a subgroup of order m. Since G' is invariant and $(m, r) = 1$, every element of order dividing m in G must be in G'. Hence all the subgroups of order m in G must be in G' where by hypothesis they form a complete set of conjugates under G' and consequently a complete set of conjugates under G.

If r divides m, then G' has a complete set of conjugates of order m/r. Since G' is invariant, the subgroups of order m/r in G' are a complete set of conjugates under G. Let M' be one of these subgroups of order m/r in G'.

If M' is the only subgroup of order m/r in G', then, it must be invariant in G. Then the quotient group is of order rg/m and since r is a Sylow divisor of the order of this quotient group, it has a subgroup of order r. Then G has a subgroup of order $r \cdot m/r$ or m. On the other hand if M' is not the only subgroup of order m/r in G', let it be one of k subgroups of order m/r in G'. Then k divides the order g/r of G' and the normalizer of M' in G is of order g/k divisible by m. Since this normalizer is a soluble group of order g/k less than g, it has a subgroup of order m.

There remains only to show, for the case r divides m, that all the subgroups of order m in G are conjugate under G. Let M be a subgroup of order m. If there is no other subgroup of order m in G, then the theorem is true by default. However, if M_1 is another subgroup

of order m in G it will be shown that M and M_1 are conjugate under G.

Let the crosscut of M and G' be Γ of order γ and the crosscut of M_1 and G' be Γ_1 of order γ_1. Then since G is generated by M and G' as well as by M_1 and G' and since G' is invariant under both M and M_1 we have

$$(g/r)(m)/\gamma = g, \qquad (g/r)(m)/\gamma_1 = g,$$

whence

$$\gamma = \gamma_1 = m/r.$$

Then since Γ and Γ_1 are of order m/r, they are conjugate under G'. If $S^{-1}\Gamma_1 S = \Gamma$, then $S^{-1}M_1S$ and M have a common invariant subgroup Γ. If Γ is invariant in G, the quotient group G/Γ is of order gr/m. Since r is a Sylow divisor, the subgroups M/Γ and $S^{-1}M_1S/\Gamma$ of order r are conjugate under G/Γ and hence M and $S^{-1}M_1S$ are conjugate under G as was to be proved. If, however, Γ is not invariant under G, its normalizer is a proper subgroup of G containing M and $S^{-1}M_1S$ which are therefore conjugate under the normalizer of Γ as was to be proved.

REED COLLEGE

ON REGULAR FAMILIES OF CURVES[1]

HASSLER WHITNEY

A family F of non-intersecting curves filling a metric space is called *regular* if, in a neighborhood of any point p, it is homeomorphic with a family of straight lines. We have given in another paper[2] a necessary and sufficient condition, which we shall call (A′) (to be described below), that a family F be regular. We shall prove in this note that the following condition is sufficient:

(A) Given any point p, and a direction on the curve through p, there is an arc pq in this direction with the following property. For every $\epsilon>0$ there is a $\delta>0$ such that for any p', with $\rho(p', p)<\delta$, there is an arc $p'q'$ of $C(p')$ such that

$$p'q' \subset V_\epsilon(pq), \qquad q' \subset V_\epsilon(q). \tag{1}$$

The condition (A′) is the same, except that after (1), we add:

(2) If r' and s' are on $p'q'$ and $\rho(r', s')<\delta$, then $\delta(r's')<\epsilon$.

From the present theorem it is clear that the families of curves recently defined by Niemitzki[3] are regular.

To prove the theorem, suppose (A) holds, but (A′) does not. Then the following is true:

(B) There is a point p, and a direction of the curve $C(p)$, such that for any arc pq on $C(p)$ in this direction, there is an $\epsilon>0$, such that for any $\delta>0$, there is a point p', with $\rho(p', p)<\delta$, such that for any q' on $C(p')$,

(3) either $p'q' \not\subset V_\epsilon(pq)$, or $q' \not\subset V_\epsilon(q)$,

---

[1] Presented to the Society, April 27, 1940.

[2] Annals of Mathematics, (2), vol. 34 (1933), pp. 244–270. We refer to this paper as RF. By RF, Theorem 7A, F is regular as there defined. The converse is proved as follows. By Theorem 17A, there is a cross-section S through p. In a neighborhood of p, the curves are orientable (this is easily seen, for instance, with the help of Theorem 9B). Choose an open subset S' of S, and let U be all points $q'=g'(q, \alpha)$, q in S', $|\alpha|<\epsilon$ (see RF, §15); U is a neighborhood of p, expressed as the product of S' and the open line segment $-\epsilon<\alpha<\epsilon$.

By a curve, we shall mean here the topological image of an open line segment or of a circle. We shall use $\rho(p, q)$ for distance, $\delta(A)$ for the diameter of the set A, and $V_\epsilon(A)$ for the set of all points p, $\rho(p, A)<\epsilon$. Let $C(p)$ mean the curve of F through p.

[3] V. Niemytzki, Recueil Mathématique de Moscou, vol. 6 (48) (1939), pp. 283–292. We mention two further papers in the subject: H. Whitney, Duke Mathematical Journal, vol. 4 (1938), pp. 222–226, showing that if the curves fill a region in 3-space, a cross-section may be chosen so as to be a 2-cell; W. Kaplan, Duke Mathematical Journal, vol. 7 (1940), pp. 154–185, studying families filling the plane.

(4) or there are points r', s' on $p'q'$ such that $\rho(r', s') < \delta$, and $\delta(r's') \geqq \epsilon$.

Choose a point p and a direction on $C = C(p)$, by (B). Choose q on C in this direction, by (A). Choose $\epsilon > 0$ by (B). For each positive integer i, choose δ_i by (A), with ϵ replaced by ϵ/i. Choose p_i by (B), with δ replaced by δ_i. Choose q_i by (A), with p' replaced by p_i; then

$$p_i q_i \subset V_{\epsilon/i}(pq), \qquad q_i \subset V_{\epsilon/i}(q). \tag{5}$$

By (B), as $\epsilon_i < \epsilon$, we may choose p_i' and q_i' on $p_i q_i$ so that

$$\rho(p_i', q_i') < \delta_i, \qquad \delta(p_i' q_i') \geqq \epsilon. \tag{6}$$

By (6), we may choose r_i on $p_i' q_i'$ so that $\rho(p_i', r_i) \geqq \epsilon/2$. By (5) and (6), we may choose a subsequence so that for some points p' and r on pq,

$$p'_{\lambda_i} \to p', \qquad q'_{\lambda_i} \to p', \qquad r_{\lambda_i} \to r; \tag{7}$$

then $r \neq p'$. Say, for definiteness, that r is in the direction of q from p'. The set of such points r which are limits of such sequences $\{r_{\lambda_i}\}$ forms a closed set, which, by (5), is in $p'q$; we shall let r be the point furthest from p'. (It might be q.)

Assuming that (A) holds for the point r and the direction away from p', we shall arrive at a contradiction. Choose a point s on C in this direction from r, by (A). (If C is a closed curve, it might happen that s is on the arc pr.) Choose r' and s' on C just behind and just in front of r, so that r' is on neither pp' nor rs, and s' is not on $p'r$. We shall show that for any $\epsilon' > 0$ there is an integer j and a point s_j on $p_j' q_j'$ within ϵ' of s'; as s is in pq, by (5), this will contradict the definition of r, and thus prove the theorem.

Set

$$4\eta = \min\,[\rho(r', rs), 2\epsilon']. \tag{8}$$

Choose r_-, r_+, s_-, s_+ on C in the order $r_- r' r_+ r s_- s' s_+ s$, so that r_- is not in rs and s_+ is not in pr (if C is closed), and so that

$$r_- r' r_+ \subset V_\eta(r'), \qquad s_- s' s_+ \subset V_\eta(s'). \tag{9}$$

Set

$$2\epsilon'' = \min\,[\rho(pr_-, r_+ q), \rho(rs_-, s_+ s), \eta]. \tag{10}$$

Using r, s, and ϵ'', choose $\delta'' > 0$ by (A). By (7), we may choose j so that

$$\epsilon/j < \epsilon'', \ \rho(p_j', p') < \epsilon'', \ \rho(q_j', p') < \epsilon'', \ \rho(r_j, r) < \delta''. \tag{11}$$

By the choice of δ'', we may choose s^* on $C(r_j)$ so that

$$r_j s^* \subset V_{\epsilon''}(rs), \qquad s^* \subset V_{\epsilon''}(s). \tag{12}$$

As $r_j s^*$ is a connected set, (11), (12) and (10) show that there is a point s_j on it such that $\rho(s_j, rs_- + s_+ s) \geqq \epsilon''$; hence, by (12), $\rho(s_j, s_- s_+) < \epsilon''$, and by (9),

$$\rho(s_j, s') < \epsilon'' + \eta < 2\eta \leqq \epsilon'. \tag{13}$$

By (12) and (8),

$$\rho(r', r_j s_j) > 2\eta. \tag{14}$$

By (5) and (11),

$$p_j' q_j' \subset V_{\epsilon''}(pq). \tag{15}$$

By (11), (10), (5) and (9), there are points p_j^* in $p_j' r_j$ and q_j^* in $r_j q_j'$ such that

$$\rho(p_j^*, r') < 2\eta, \qquad \rho(q_j^*, r') < 2\eta. \tag{16}$$

By this and (14), the arc $r_j s_j$ is contained in the arc $p_j^* r_j q_j^* \subset p_j' q_j'$. Hence $s_j \subset p_j' q_j'$, which, with (13), gives the contradiction.

HARVARD UNIVERSITY

TRANSLATED FUNCTIONS AND STATISTICAL INDEPENDENCE[1]

R. P. AGNEW AND M. KAC

1. **Introduction.** Professor Steinhaus[2] proposed the question whether there exists a function $f(t)$ defined in $(-\infty, \infty)$ such that, for each sequence $\lambda_1, \lambda_2, \cdots$ of different real numbers, the "translated" functions $f(t+\lambda_1), f(t+\lambda_2), \cdots$ form a sequence of statistically independent functions.[3] We shall answer this question in the affirmative by giving concrete examples, and shall discuss some related problems.

2. **Notation and lemmas.** Let

$$\mathfrak{M}\{f(t)\} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(t)dt \tag{1}$$

denote as usual the *mean value* of $f(t)$ in case the limit exists. If E is a point set on the line $(-\infty, \infty)$ and $g(t)$ is the characteristic function of E, the mean value $\mathfrak{M}\{g(t)\}$ (if it exists) is called the *relative measure* of E and will be denoted by $|E|$. A non-decreasing function $\sigma(\alpha)$ such that $\sigma(-\infty)=0$, $\sigma(+\infty)=1$ is called an *asymptotic distribution function* of $f(t)$ if

$$|E_t\{f(t) < \alpha\}| = \sigma(\alpha) \tag{2}$$

at each point of continuity of $\sigma(\alpha)$. A set $f_1(t), \cdots, f_n(t)$ of functions having asymptotic distribution functions $\sigma_1(\alpha), \cdots, \sigma_n(\alpha)$ is called *statistically independent* if

$$|E_t\{f_1(t) < \alpha_1; \cdots; f_n(t) < \alpha_n\}| = \sigma_1(\alpha_1)\cdots\sigma_n(\alpha_n) \tag{3}$$

for each set $\alpha_1, \cdots, \alpha_n$ of real numbers such that α_k is a point of continuity of $\sigma_k(\alpha)$. An infinite set of functions is called statistically independent if each finite subset is statistically independent.

Our proofs will be based on the following theorem of Kac and Steinhaus, loc. cit.

---

[1] Presented to the Society, February 24, 1940.

[2] In a letter to one of the authors. The problem arose in connection with the theory of turbulence, but because of the outbreak of the war we have not been able to get details.

[3] See M. Kac and H. Steinhaus, *Sur les fonctions indépendantes* IV, Studia Mathematica, vol. 7 (1937), pp. 1–15; and P. Hartman, E. R. van Kampen, and A. Wintner, *Asymptotic distributions and statistical independence*, American Journal of Mathematics, vol. 61 (1939), pp. 477–486.

LEMMA 1. *A set $f_1(t), \cdots, f_n(t)$ of bounded functions having absolutely continuous distribution functions is statistically independent if and only if the equality*

$$\mathfrak{M}\left\{\prod_{j=1}^{n} f_j^{k_j}(t)\right\} = \prod_{j=1}^{n} \mathfrak{M}\{f_j^{k_j}(t)\} \tag{4}$$

holds for each set $k_1, k_2, \cdots, k_n$ of positive integer exponents.

Repeated use will be made of the following simple lemma.

LEMMA 2. *If $\phi(t)$ is a real differentiable function and constants $\tau>0$ and $\alpha>0$ exist such that $\phi'(t)$ is monotone and $|\phi'(t)|>\alpha$ over each of the intervals $(-\infty, -\tau)$ and (τ, ∞), then*

$$\mathfrak{M}\{e^{i\phi(t)}\} = 0. \tag{5}$$

Let I_1, I_2, and I_3 denote the integrals of $\exp i\phi(t)$ over $(-T, -\tau)$, $(-\tau, \tau)$, and (τ, T) respectively. Applying the second mean value theorem we obtain, when $T>\tau$,

$$\begin{aligned} I_3 &= \int_\tau^T \frac{\phi'(t)\cos\phi(t)}{\phi'(t)}\,dt + i\int_\tau^T \frac{\phi'(t)\sin\phi(t)}{\phi'(t)}\,dt \\ &= \frac{1}{\phi'(\tau)}\int_\tau^{\xi_1} \phi'(t)\cos\phi(t)\,dt + \frac{1}{\phi'(T)}\int_{\xi_1}^T \phi'(t)\cos\phi(t)\,dt \\ &\quad + \frac{i}{\phi'(\tau)}\int_\tau^{\xi_2} \phi'(t)\sin\phi(t)\,dt + \frac{i}{\phi'(T)}\int_{\xi_2}^T \phi'(t)\sin\phi(t)\,dt, \end{aligned}$$

and evaluating the last four integrals gives $|I_3|<8/\alpha$. Likewise $|I_1|<8/\alpha$, and since $|I_2|\leqq 2\tau$ our lemma follows.

If $\phi(t)$ is a real polynomial in t of positive degree, Lemma 2 is obviously applicable. The same is true if $\phi(t)=\sum_{j=1}^{n} a_j \exp(t+\lambda_j)^2$ where $\lambda_1, \cdots, \lambda_n$ are different real numbers and $a_1, \cdots, a_n$ are real coefficients not all 0.

3. **An example.** We can now establish our main result.

THEOREM 1. *The translations of the function*

$$f(t) = \cos e^{t^2} \tag{6}$$

form a statistically independent set; that is, for each set $\lambda_1, \cdots, \lambda_n$ of different real numbers, the functions $f(t+\lambda_1), \cdots, f(t+\lambda_n)$ are statistically independent.

Our first step is to verify that (4) holds. We find that

$$\prod_{j=1}^{n} \cos^{k_j} e^{(t+\lambda_j)^2} = \prod_{j=1}^{n} 2^{-k_j}[\exp(ie^{(t+\lambda_j)^2}) + \exp(-ie^{(t+\lambda_j)^2})]^{k_j}$$

$$= \prod_{j=1}^{n} 2^{-k_j} \sum_{\alpha_j=0}^{k_j} C_{k_j,\alpha_j}[\exp(i\alpha_j e^{(t+\lambda_j)^2})]$$

$$\cdot [\exp(-i(k_j-\alpha_j)e^{(t+\lambda_j)^2})]$$

$$= \prod_{j=1}^{n} 2^{-k_j} \sum_{\alpha_j=0}^{k_j} C_{k_j,\alpha_j} \exp i(2\alpha_j - k_j)e^{(t+\lambda_j)^2}$$

$$= 2^{-k_1} \cdots 2^{-k_n} \sum_{\alpha_1=0}^{k_1} \cdots \sum_{\alpha_n=0}^{k_n} C_{k_1,\alpha_1} \cdots C_{k_n,\alpha_n}$$

$$\cdot \exp i \sum_{j=1}^{n} (2\alpha_j - k_j)e^{(t+\lambda_j)^2}.$$

Since $\mathfrak{M}\{f\}$ is an additive homogeneous operation, we can apply Lemma 2 to show that

$$\mathfrak{M}\left\{\prod_{j=1}^{n} \cos^{k_j} e^{(t+\lambda_j)^2}\right\} = 0 \tag{7}$$

if at least one of $k_1, k_2, \cdots, k_n$ is odd; but

$$\mathfrak{M}\left\{\prod_{j=1}^{n} \cos^{k_j} e^{(t+\lambda_j)^2}\right\} = \prod_{j=1}^{n} 2^{-k_j} C_{k_j,k_j/2} \tag{8}$$

if $k_1, k_2, \cdots, k_n$ are all even. Noticing that the particular case $n=1$ gives

$$\begin{aligned} \mathfrak{M}\{\cos^k e^{(t+\lambda)^2}\} &= 0, && k \text{ odd}, \\ &= 2^{-k} C_{k,k/2}, && k \text{ even}, \end{aligned} \tag{9}$$

we see that the mean value of the product is the product of the mean values as in (4). Moreover the fact that the moments of $f(t) \equiv \cos \exp (t+\lambda)^2$ are the same as the moments of $\cos t$ implies[4] that the distribution function of $f(t)$ is the same as that of $\cos t$ and hence is absolutely continuous. Therefore we can apply Lemma 1 to obtain the conclusion of Theorem 1.

4. **Functions which oscillate less rapidly.** The function $\cos \exp t^2$ of Theorem 1 oscillates very rapidly as $|t| \to \infty$. A natural question

[4] A. Wintner, *Über die statistiche Unabhängigkeit der asymptotischen Verteilungsfunktionen inkommensurabler Partialschwingungen*, Mathematische Zeitschrift, vol. 36 (1933), pp. 618–629.

arises whether $\exp t^2$ can be replaced by a positive integer power of t, say t^m. If we put $f(t)=\cos t^m$, then the answer to the question of statistical independence of the translations $f(t+\lambda_1), f(t+\lambda_2), \cdots$ is only partially positive.

THEOREM 2. *If $f(t)=\cos t^m$, m being a positive integer, then each set of m different translations $f(t+\lambda_1), \cdots, f(t+\lambda_m)$ is a statistically independent set; but $(m+1)$ different numbers $\lambda_1, \cdots, \lambda_{m+1}$ can be chosen in such a way that the corresponding set of $(m+1)$ translations is not statistically independent.*

Proof of this theorem is similar to that of Theorem 1. For each $n=1, 2, \cdots$ we find on setting

$$P_n(t) = \prod_{j=1}^{n} \cos^{k_j} (t+\lambda_j)^m \tag{10}$$

that

$$\mathfrak{M}\{P_n(t)\} = 2^{-k_1} \cdots 2^{-k_n} \sum_{\alpha_1=0}^{k_1} \cdots \sum_{\alpha_n=0}^{k_n} C_{k_1,\alpha_1} \cdots C_{k_n,\alpha_n} \mathfrak{M}\{e^{i\phi(t)}\} \tag{11}$$

where $\phi(t) \equiv \phi(\lambda_1, \cdots, \lambda_n; k_1, \cdots, k_n; \alpha_1, \cdots, \alpha_n; t)$ is given by

$$\phi(t) = \sum_{j=1}^{n} (2\alpha_j - k_j)(t+\lambda_j)^m = \sum_{s=0}^{m} \left[C_{m,s} \sum_{j=1}^{n} (2\alpha_j - k_j)\lambda_j^s \right] t^{m-s}. \tag{12}$$

Applying Lemma 2, we see that $\mathfrak{M}\{\exp i\phi(t)\} \neq 0$ if and only if

$$\sum_{j=1}^{n} (2\alpha_j - k_j)\lambda_j^s = 0, \quad s = 0, 1, \cdots, m-1. \tag{13}$$

In case $n=m$, (13) is a system of m homogeneous linear equations in m quantities $(2\alpha_j - k_j)$ whose determinant (Vandermonde's determinant) does not vanish; and if (13) holds, then $2\alpha_j = k_j$ when $j=1, 2, \cdots, m$. Hence, in case $n=m$, $\mathfrak{M}\{\exp i\phi(t)\} = 0$ unless $k_1, \cdots, k_m$ are even and $\alpha_j = k_j/2$ when $j=1, 2, \cdots, m$. Therefore, in case $n=m$, formulas analogous to (7) and (8) hold and statistical independence of $f(t+\lambda_1), \cdots, f(t+\lambda_m)$ follows as in the proof of Theorem 1.

In case $m>1$ and $\lambda_1, \lambda_2, \cdots, \lambda_n$ are linearly independent (in the sense that no one of $\lambda_1, \cdots, \lambda_n$ is a linear combination with rational coefficients of the remaining) then (13) again implies that $2\alpha_j = k_j$ and we conclude that $f(t+\lambda_1), \cdots, f(t+\lambda_n)$ are statistically independent. In case $m>1$, this shows existence of infinite sets of statistically independent translations of $\cos t^m$.

We complete the proof of Theorem 2 by proving the following theorem.

THEOREM 3. *If m is a positive integer and $\lambda_1, \lambda_2, \cdots, \lambda_{m+1}$ are $m+1$ different rational numbers, then the set of $(m+1)$ translations*

$$\cos(t+\lambda_1)^m, \cos(t+\lambda_2)^m, \cdots, \cos(t+\lambda_{m+1})^m \tag{14}$$

is not statistically independent.

We use the formulas obtained by setting $n=m+1$ in (10), (11), (12), and (13). The system

$$\sum_{j=1}^{m+1} \lambda_j^s x_j = 0, \qquad s = 0, 1, \cdots, m-1, \tag{15}$$

of m equations in $(m+1)$ unknowns $x_1, \cdots, x_{m+1}$ is such that if a nonzero rational value is assigned to one of the x's, then the remaining x's are uniquely determined and are rational. It follows that (15) has non-trivial solutions in integers. Let X denote the set of points $x \equiv (x_1, \cdots, x_{m+1})$ in $(m+1)$-dimensional space for which $x_1, \cdots, x_{m+1}$ are integers, not all 0, satisfying (15). For each $x \in X$, let $\|x\|^2 = x_1^2 + \cdots + x_{m+1}^2$. Since $\|x\|^2$ is necessarily a positive integer, the minimum value of $\|x\|^2$ for $x \in X$ is an integer $R>0$. Choose $y \equiv (y_1, \cdots, y_{m+1}) \in X$ such that $\|y\|^2 = R$, and let $k_1, \cdots, k_{m+1}$ be defined by

$$k_j = |y_j|, \qquad j = 1, 2, \cdots, m+1. \tag{16}$$

If $0 \leqq \alpha_j \leqq k_j$ for each $j = 1, 2, \cdots, m+1$ while $0 < \alpha_j < k_j$ for at least one j, then

$$\sum_{j=1}^{m+1} (2\alpha_j - k_j)^2 < \sum_{j=1}^{m+1} k_j^2 = \sum_{j=1}^{m+1} y_j^2 = \|y\|^2 = R \tag{17}$$

and the minimal property of R implies that the point

$$(2\alpha_1 - k_1, 2\alpha_2 - k_2, \cdots, 2\alpha_{n+1} - k_{n+1}) \tag{18}$$

is not in the set X; this means that at least one of the equations in (13) fails to hold, and accordingly $\mathfrak{M}\{\exp i\phi(t)\} = 0$. The remaining 2^{m+1} terms in the sum (11) are those for which each α_j is either 0 or k_j and accordingly (18) is a point on the sphere (hypersphere) S with center at $(0, \cdots, 0)$ and radius $R^{1/2}$. But the set X, being on each of m independent planes (hyperplanes), is on their line of intersection; hence at most two of the points on S can be in the set X. It follows that the point (18) fails to be in X, and hence that $\mathfrak{M}\{\exp i\phi(t)\} = 0$,

unless the indices $\alpha_1, \cdots, \alpha_{m+1}$ are such that (i) $2\alpha_j - k_j = y_j$ for each $j=1, \cdots, m+1$ or (ii) $2\alpha_j - k_j = -y_j$ for each $j=1, \cdots, m+1$. In case (i), (13) holds so that $\phi(t)$ is independent of t, and we see that $\phi(t)=\theta$ where

$$\theta = \sum_{j=1}^{m+1} y_j \lambda_j^m; \tag{19}$$

in case (ii), $\phi(t) = -\theta$. Therefore

$$\begin{aligned} \mathfrak{M}\left\{ \prod_{j=1}^{m+1} \cos^{k_j} (t+\lambda_j)^m \right\} &= 2^{-k_1} \cdots 2^{-k_{m+1}} (\mathfrak{M}\{e^{i\theta}\} + \mathfrak{M}\{e^{-i\theta}\}) \\ &= 2^{-k_1} \cdots 2^{-k_{m+1}} (2\cos\theta) \neq 0, \end{aligned} \tag{20}$$

the last step being a consequence of the fact that the right member of (19) is rational and hence that θ cannot be an odd multiple of $\pi/2$.

At least one of $k_1, k_2, \cdots, k_{n+1}$ must be odd, for otherwise $z \equiv (y_1/2, \cdots, y_{m+1}/2)$ would be a point of X with $\|z\|^2 = \|y\|^2/4 = R/4 < R$; therefore

$$\prod_{j=1}^{m+1} \mathfrak{M}\{\cos^{k_j} (t+\lambda_j)^m\} = 0. \tag{21}$$

Comparing (20) and (21), we see that (4) fails when $n=m+1$ and $f_j(t) = \cos(t+\lambda_j)^m$; hence Lemma 1 implies that the $(m+1)$ translations in (14) are not statistically independent. This completes proof of Theorems 3 and 2.

Anyone who wishes to check Theorem 3 by consideration of a simple special case will find that if $m=2$, $\lambda_1=1$, $\lambda_2=2$, $\lambda_3=3$ then $y \equiv (y_1, y_2, y_3)$ may be either $(1, -2, 1)$ or $(-1, 2, -1)$; that $k_1=1$, $k_2=2$, $k_3=1$; and that elementary formulas for products of cosines and use of Lemma 2 gives

$$\mathfrak{M}\{\cos(t+1)^2 \cos^2(t+2)^2 \cos(t+3)^2\} = 8^{-1} \cos 2, \tag{22}$$

which is a special case of (20).

5. **Other functions.** Our methods can be used to show that certain rapidly oscillating functions $f(t)$ other than the function $\cos \exp t^2$ of §3 have statistically independent translations. Our first solution of the problem of §1 was a "direct" (but too tedious to be given here) proof of the fact that the set of all translations of the function $\cos \exp \exp t^2$ is statistically independent. Our "direct" method uses the definition of statistical independence rather than the criterion of Lemma 1, and applies only to functions which oscillate very rapidly.

To handle such functions as $\cos t^m$, where the question of statistical independence of translations is more delicate, we need the criterion of Lemma 1.

6. **Applications.** As soon as statistical independence of a set of functions is established, we can apply the general theory of addition of independent random variables.[5] As an illustration, we mention the following result: If $f(t) = \cos \exp t^2$, and $\lambda_1, \lambda_2, \cdots$ is a sequence of different real numbers, then

$$(23) \quad \lim_{n\to\infty} \left| E_t\{f(t+\lambda_1) + \cdots + f(t+\lambda_n) < \alpha n^{1/2}\} \right| = \frac{1}{\pi^{1/2}} \int_{-\infty}^{\alpha} e^{-u^2} du.$$

CORNELL UNIVERSITY

[5] See Kac and Steinhaus, loc. cit., and further references given there. Also P. Lévy, *Théorie de l'Addition des Variables Aléatoires*, Paris, Gauthier-Villars, 1937.

NOTE ON A CERTAIN TYPE OF DIOPHANTINE SYSTEM[1]

E. T. BELL

1. **The quadratic case.** Several writers[2] have considered the problem of making a^2x^2+dx, b^2x^2+ex, $\cdots$ simultaneously squares, a, d, b, e, $\cdots$ being constants, without discussing the conditions under which an integer solution exists. The fact that the coefficients of x^2 are squares is of no consequence. It is shown in §2 how necessary and sufficient conditions for the existence of integer solutions in much more general problems (§2, (9), (13)) can be determined and how, when these conditions are satisfied, the solutions may be found. The details are given first for three very special cases, (1), (5), (6) below.

If all letters denote integers, and a, b, c, d are constant, we seek necessary and sufficient conditions that the system

$$ax^2 + bx = y^2, \qquad cx^2 + dx = z^2 \tag{1}$$

shall have a solution x, y, z. We shall assume that $abcd \neq 0$, as the excluded possibilities require only slight modifications, all of which are included in the general method.

The required conditions are that b, d be simultaneously representable in two forms of degree seven. Precisely, (1) has a solution in integers x, y, z if and only if integers f, g, h, k, m, n exist such that

$$b = hf^2(m^2 - ag^2k^2), \qquad d = hg^2(n^2 - cf^2k^2). \tag{2}$$

Provided such f, g, $\cdots$ exist, the complete solution of (1) is

$$x = hf^2g^2k^2, \qquad y = hf^2gkm, \qquad z = hfg^2kn,$$

where f, g, h, k, m, n run through all solutions of (2).

To prove this, we rewrite (1) as

$$x(ax + b) = y^2, \qquad x(cx + d) = z^2,$$

which is of the form

$$xu = y^2, \qquad xv = z^2, \tag{3}$$

with

$$u \equiv ax + b, \qquad v \equiv cx + d.$$

The complete integer solutions of the respective equations in (3),

[1] Presented to the Society, April 6, 1940.

[2] Summary of results to 1920 in L. E. Dickson, *History of the Theory of Numbers*, vol. 2, 1920, chap. 18.

with x, y, z, u, v unrestricted variables, are

$$x = r_1 s_1^2, \qquad u = r_1 t_1^2, \qquad y = r_1 s_1 t_1;$$
$$x = r_2 s_2^2, \qquad v = r_2 t_2^2, \qquad z = r_2 s_2 t_2,$$

in which all the letters with suffixes denote arbitrary integers. Since in the system (3) the two values of x must be equal, we have

$$r_1 s_1^2 = r_2 s_2^2, \tag{4}$$

of which the complete integer solution is

$$r_1 = hf^2, \qquad s_1 = kg, \qquad r_2 = hg^2, \qquad s_2 = kf,$$

where f, g, h, k are arbitrary integers. From the (restricted) values $u \equiv ax+b$, $v \equiv cx+d$ of u, v, we find the stated conditions on b, d, after the change of notation $(t_1, t_2) \to (m, n)$. The completeness of the solutions of (3), (4) implies the completeness of the solution of (1) when $f, g, \cdots$ run through all solutions of (2). This treatment is possible only because (3), equivalent to (1), is a pure multiplicative system. A full discussion of such systems, with a non-tentative method for finding their complete solutions, was given in a former paper.[3]

The above solutions of (3), (4) were obtained by this method.

Proceeding similarly with

$$ax^2 + bx = y^2, \qquad cx^2 + dx = z^2, \qquad ex^2 + jx = w^2, \tag{5}$$

in which a, b, c, d, e, j are constant integers, different from zero (not an essential restriction), we find that all solutions x, y, z, w are given by

$$x = hf^2g^2k^2p^2q^2r^2s^2, \qquad y = hf^2gkpq^2rs^2l,$$
$$z = hf^2g^2k^2pqrsm, \qquad w = hfg^2kpq^2r^2sn,$$

where $h, \cdots, m$ run through all integer solutions of

$$b = hf^2q^2s^2(l^2 - ak^2g^2p^2r^2), \qquad d = hf^2g^2k^2(m^2 - cp^2q^2r^2s^2),$$
$$j = hg^2q^2r^2(n^2 - ek^2f^2p^2s^2).$$

Here the simultaneous representation of b, d, j which is necessary and sufficient for the existence of a solution of (5) is by forms of degree fifteen. This may suggest that anything approaching a complete solution of the systems quoted[2] is not to be expected in the near future. However, any number of solvable systems can be constructed immedi-

[3] American Journal of Mathematics, vol. 55 (1933), pp. 50–66.

ately from the conditions, and likewise for such systems having infinities of solutions.

The pure multiplicative system corresponding to (3) for (5) is

$$r_1 s_1^2 = r_2 s_2^2 = r_3 s_3^2,$$

of which the complete solution is

$$r_1 = hf^2q^2s^2, \qquad r_2 = hf^2g^2k^2, \qquad r_3 = hg^2q^2r^2,$$
$$s_1 = kgpr, \qquad s_2 = pqrs, \qquad s_3 = kfps.$$

The special features of (1), (5) and the systems cited[2] which make it possible to treat the equations as a pure multiplicative system are that each equation lacks a constant term in one of its members and that its other member consists of a single term. The second of these restrictions is removed in §2. It is immaterial that the single term has the coefficient unity; for multiplicative systems with arbitrary constant integer coefficients can also be solved completely by a straightforward, non-tentative process.[3]

For example, the pure multiplicative system corresponding to the system

$$ax^2 + bx = ry^2, \qquad cx^2 + dx = sz^2, \tag{6}$$

in which the constant integers a, b, c, d, r, s are all different from zero, is

$$xu = ry^2, \qquad xv = sz^2. \tag{7}$$

We first solve (7) as if all the letters were variable integers, and find the complete solution

$$x = tr_1^2 r_2^2 r_4 s_1 s_2^2 s_4^2 x_1 x_2^2 y_1^2 y_2^2, \qquad y = r_2 s_1 s_2 s_4 x_1 x_2 y_1^2 y_2 z_1,$$
$$z = r_1 r_2 r_4 s_2 x_1 x_2 y_1^2 y_2 z_2, \tag{8}$$
$$u = r_2 r_3 s_1 x_1 y_1^2 z_1^2, \qquad v = r_1 s_2 s_3 x_1 y_2^2 z_2^2,$$
$$r = tr_1 r_2 r_3 r_4^2, \qquad s = ts_1 s_2 s_3 s_4^2,$$

in which t and all the letters with suffixes denote arbitrary integers. If now r, s are constant, we resolve them in all possible ways into products of five factors as indicated above, and it suffices (as appears in the method cited[2]) to take t equal to the G.C.D. of r, s. With each such pair of resolutions of r, s there is associated a solution x, y, z, u, v as given in (8); all the solutions of (7), with r, s constant, fall into sets

determined by these pairs. Proceeding as in (1)–(4), we find that (6) has an integer solution x, y, z if and only if integers $x_1, x_2, y_1, y_2, z_1, z_2$ exist such that, for some resolution of r, s as in (8),

$$b = r_2 s_1 x_1 y_1^2 (r_2 z_1^2 - a t r_1 r_2^2 r_4^2 s_2^2 s_4^2 x_2^2 y_2^2),$$
$$d = r_1 s_2 x_1 y_2^2 (s_2 z_2^2 - c t r_2^2 r_4^2 s_1^2 s_2^2 s_4^2 x_2^2 y_1^2);$$

and if these conditions are satisfied, the complete solution x, y, z of (7) is as given in (8), in which t and the r_i, s_i refer to all resolutions of the constants r, s of the type indicated. The extension to a system of n (>2) quadratics is immediate.

2. **The general case.** The notation is as follows.

$P_i(x)$ is a polynomial in x of degree m_i+r_i, $m_i>0$, $r_i>0$, with constant integer coefficients.

The term of lowest degree in $P_i(x)$ is $a_i x^{m_i}$, $a_i \neq 0$.

$$P_i(x) \equiv x^{m_i} Q_i(x), \qquad Q_i(x) \equiv R_i(x) + a_i,$$

so that $R_i(x)$ is a polynomial in x of degree $r_i>0$ with constant integer coefficients and no constant term.

b_i, n_i, s denote constant integers, $b_i \neq 0$, $n_i>0$, $s>1$.

By the method exemplified in §1, we may obtain necessary and sufficient conditions that the system

$$P_i(x) = b_i y_i^{n_i}, \qquad i = 1, \cdots, s, \tag{9}$$

shall have a solution in integers $x, y_1, \cdots, y_s$. The conditions concern the representability of the constant integers $a_1, \cdots, a_s$ in certain forms; the complete solution of (9) is known when all these representations of $a_1, \cdots, a_s$ are known.

Rewriting (9) as

$$x^{m_i}[R_i(x) + a_i] = b_i y_i^{n_i}, \qquad i = 1, \cdots, s, \tag{10}$$

we associate with (10) the pure multiplicative system

$$x^{m_i} z_i = b_i y^{n_i}, \qquad i = 1, \cdots, s. \tag{11}$$

The complete integer solution x, z_i, y_i, $i=1, \cdots, s$, of (11) is found as in §1 in terms of t, $\equiv t(m_1, \cdots, m_s, n_1, \cdots, n_s)$, parameters; say the solution is

$$x = f(u_1, \cdots, u_t), \quad z_i = g_i(u_1, \cdots, u_t), \quad y_i = h_i(u_1, \cdots, u_t), \quad i = 1, \cdots, s. \tag{12}$$

The f, g_i, h_i are power products in $u_1, \cdots, u_t$ with integer coefficients which are power products formed from the factors in certain resolutions of the constants $b_1, \cdots, b_s$ into products; the forms of f, g_i, h_i and the coefficients described depend only on the given constant integers m_i, n_i. Necessary and sufficient conditions that the system (9) have a solution in integers x, y_i, $i=1, \cdots, s$, are found as in §1, and refer to the representation of the constants a_i in certain forms. Comparing (10), (11), we find

$$a_i = g_i(u_1, \cdots, u_t) - R_i(f(u_1, \cdots, u_t)), \qquad u = 1, \cdots, s,$$

as the required conditions. When these are satisfied, the solutions are obtained from (12).

An obvious modification gives necessary and sufficient conditions for the existence of an integer solution, and so on, of the more general system

$$c_iP_i(x) = b_iT_i(y_i), \qquad i = 1, \cdots, s, \tag{13}$$

in which c_i, b_i are constants, $P_i(x)$ is as before, and $T_i(y_i)$ is a polynomial in y_i of degree n_i+t_i, $n_i>0$, $t_i>0$, in which the term of lowest degree is $d_iy_i^{t_i}$, $d_i\neq 0$. The associated pure multiplicative system is

$$c_ix^{m_i}z_i = b_iy_i^{n_i}w_i, \qquad i = 1, \cdots, s;$$

the conditions concern the representability of the a_i, d_i in certain forms. As in all cases the conditions are both necessary and sufficient, the complete solution of the given system is equivalent to finding the total representation of the coefficients a_i, d_i in a determinable system of forms.

California Institute of Technology

POWER SERIES THE ROOTS OF WHOSE PARTIAL SUMS LIE IN A SECTOR[1]

LOUIS WEISNER

If the roots of the partial sums of a power series $f(z)=\sum a_n z^n$ lie in a sector with vertex at the origin and aperture $\alpha<2\pi$, the power series cannot have a positive finite radius of convergence.[2] But if $f(z)$ is an entire function, the roots of its partial sums may lie in such a sector. The question arises: what restrictions are imposed on $f(z)$ by the requirement that α be sufficiently small, say $\alpha<\pi$? According to a theorem of Pólya the order of $f(z)$ must be not greater than 1 if the radius of convergence of the power series is positive.[3] Without this assumption the investigation which follows shows that *if* $\alpha<\pi$, $f(z)$ *is an entire function of order* 0. This result was obtained by Pólya for the case in which $\alpha=0$.[4]

LEMMA. *If the complex numbers* $z_1, \cdots, z_n$ $(z_1 \cdots z_n \neq 0)$ *lie in a sector with vertex at the origin and aperture* $\alpha<\pi$, *then*

$$(1) \qquad \frac{n\cos\alpha/2}{\left|\sum_{k=1}^{n} z_k^{-1}\right|} \leqq \left|z_1 \cdots z_n\right|^{1/n} \leqq \frac{1}{n}\sec\alpha/2\left|\sum_{k=1}^{n} z_k\right|.$$

When $\alpha=0$ *equality occurs if and only if* $z_1=\cdots=z_n$. *When* $\alpha>0$ *equality occurs if and only if* n *is even and* $n/2$ *of the numbers are equal to* $re^{i\phi}$ $(r>0;\ 0\leqq\phi<2\pi)$ *and the other* $n/2$ *numbers are equal to* $re^{i(\phi+\alpha)}$.

Suppose first that the sector is $-\alpha/2\leqq \operatorname{am} z\leqq\alpha/2$. Let the n numbers be

$$z_k=\left|z_k\right|e^{i\theta_k}, \qquad -\alpha/2\leqq\theta_k\leqq\alpha/2;\ k=1,\cdots,n.$$

Since

$$(2) \qquad \sum_{k=1}^{n} z_k=\sum_{k=1}^{n}\left|z_k\right|\cos\theta_k+i\sum_{k=1}^{n}\left|z_k\right|\sin\theta_k,$$

---

[1] Presented to the Society, April 27, 1940.

[2] This follows from Jentzsch's theorem: every point on the circle of convergence of a power series is a limit point of roots of its partial sums. See R. Jentzsch, *Untersuchungen zur Theorie der Folgen analytischer Funktionen*, Acta Mathematica, vol. 41 (1917), p. 219; E. C. Titchmarsh, *Theory of Functions*, 1932, p. 238.

[3] G. Pólya, *Ueber Annäherung durch Polynome deren sämtliche Wurzeln in einen Winkelraum fallen*, Nachrichten der Gesellschaft der Wissenschaften zu Göttingen, 1913, pp. 325–330.

[4] G. Pólya, *Ueber Annäherung durch Polynome mit lauter reellen Wurzeln*, Rendiconti del Circolo Matematico di Palermo, vol. 36 (1913), pp. 279–295.

$$\left|\sum_{k=1}^{n} z_k\right| \geqq \sum_{k=1}^{n} |z_k| \cos\theta_k \geqq \cos\alpha/2 \sum_{k=1}^{n} |z_k|. \tag{3}$$

Now

$$\frac{1}{n}\sum_{k=1}^{n} |z_k| \geqq |z_1 \cdots z_n|^{1/n}. \tag{4}$$

Consequently

$$\frac{1}{n}\left|\sum_{k=1}^{n} z_k\right| \geqq \cos\alpha/2\, |z_1 \cdots z_n|^{1/n}.$$

Since the numbers $z_1^{-1}, \cdots, z_n^{-1}$ also lie in the sector $-\alpha/2 \leqq \text{am } z \leqq \alpha/2$, we have

$$\frac{1}{n}\left|\sum_{k=1}^{n} z_k^{-1}\right| \geqq \cos\alpha/2\, |z_1^{-1} \cdots z_n^{-1}|^{1/n}.$$

Combining the last two inequalities, (1) results.

When $\alpha = 0$, (1) reduces to the well known relation among the harmonic, geometric and arithmetic means of n positive numbers. Here equality occurs if and only if $z_1 = \cdots = z_n$.

If equality occurs in (1) it also occurs in (3) and (4). By (4), $|z_1| = \cdots = |z_n|$. By (3), $\cos\theta_k = \cos\alpha/2$; hence $\theta_k = \pm\alpha/2$ $(k = 1, \cdots, n)$. By (2), $\sum_{k=1}^{n} \sin\theta_k = 0$. Therefore if $\alpha > 0$, n must be even, and $n/2$ of the numbers equal $re^{-i\alpha/2}$, while the other $n/2$ numbers equal $re^{i\alpha/2}$. Conversely, when these conditions are satisfied, equality is attained in (1).

If the numbers are in the sector $\phi \leqq \text{am } z \leqq \text{am } (\phi + \alpha)$, we apply the transformation

$$z' = e^{-i(\alpha/2+\phi)} z,$$

which rotates this sector into the sector $-\alpha/2 \leqq \text{am } z \leqq \alpha/2$ without affecting the value of any member of (1).

THEOREM. *If, for each $n \geqq n_0$, the roots of the partial sum of degree n of the formal power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ lie in some sector with vertex at the origin and aperture $\alpha < \pi$,*[5] *then $f(z)$ is an entire function of order 0.*

The case in which $f(z)$ is a polynomial is trivial and is excluded from

[5] While α is independent of n, we do not require that there shall be one sector which contains the roots of all the partial sums of degree $n \geqq n_0$; the lines bounding the sector may be different for different values of n.

consideration. We shall suppose $a_0 \neq 0$; otherwise a power of z could be removed from $f(z)$ without affecting the theorem. Let

$$f_n(z) = \sum_{k=0}^{n} a_k z^k, \qquad n \geqq n_0.$$

By the Gauss-Lucas theorem the roots of $f_n'(z)$ are also in the sector which contains the roots of $f_n(z)$, and the only roots of $f_n'(z)$ that lie on the boundary of the sector are multiple roots of $f_n(z)$; hence $f_n'(0) \neq 0$. Repeated applications of this argument yield the result that $a_k \neq 0$ $(k = 0, 1, \cdots)$.

According to the lemma, if $z_1, \cdots, z_n$ denote the zeros of $f_n(z)$,

$$(5) \qquad nc \left| \frac{a_0}{a_1} \right| \leqq \left| \frac{a_0}{a_n} \right|^{1/n} \leqq \frac{1}{nc} \left| \frac{a_{n-1}}{a_n} \right|, \qquad c = \cos \alpha/2; \; n \geqq n_0.$$

From the first two members of this inequality it follows that $|a_n|^{-1/n} \to \infty$ with n. Therefore $f(z)$ is an entire function. If ρ is its order,

$$(6) \qquad \frac{1}{\rho} = \liminf_{n \to \infty} \frac{\log |a_n|^{-1}}{n \log n}.$$

From the last two members of (5) we have

$$(7) \qquad \frac{1}{n} \log \frac{1}{|a_n|} - \frac{1}{n-1} \log \frac{1}{|a_{n-1}|} \geqq \frac{1}{n(n-1)} \log |a_0| + \frac{1}{n-1} \log nc.$$

Let $m = \max(n_0, 4)$, $n > m$. Substituting $n = m, m+1, \cdots, n$ in (7), and adding, we obtain

$$\frac{1}{n} \log \frac{1}{|a_n|} \geqq \frac{1}{m-1} \log \frac{1}{|a_{m-1}|} + \log |a_0| \sum_{s=m}^{n} \frac{1}{s(s-1)} + \sum_{s=m}^{n} \frac{\log s}{s-1} + \log c \sum_{s=m}^{n} \frac{1}{s-1}.$$

Now

$$\sum_{s=m}^{n} \frac{\log s}{s-1} > \sum_{s=m}^{n} \frac{\log (s-1)}{s-1} > \frac{1}{2} \log^2 n - \frac{1}{2} \log^2 (m-1),$$

$$\sum_{s=m}^{n} \frac{1}{s-1} > \log n - \log (m-1).$$

Consequently

$$\frac{1}{n}\log\frac{1}{|a_n|} > A + \frac{1}{2}\log^2 n + \log c \log n, \tag{8}$$

where A is bounded as $n\to\infty$. Comparing (6) and (8), we conclude that $\rho=0$.

HUNTER COLLEGE OF THE CITY OF NEW YORK

ON A RESULT OF HUA FOR CUBIC POLYNOMIALS[1]

ALVIN SUGAR

In a paper by L. K. Hua,[2] we find the following principal result.

THEOREM. *For any positive integer ϵ, every integer can be expressed in infinitely many ways as a sum of seven values of the cubic function*

$$f(x) = \epsilon(x^3 - x)/6 + x$$

for integral values of x; also, every integer can be expressed in infinitely many ways as a sum of seven values of

$$F(x) = (x^3 - x)/6$$

for integral values of x.[3]

In this paper we get a better result by applying a known identity for cubes to generalizations of the above polynomials. We state our results in the following two theorems. Note that, in Theorem 1, ϵ may be positive or negative, or, for that matter, zero.

Unless otherwise stated all letters in this paper stand for integers, positive, negative, or zero.

THEOREM 1. *For any ϵ, c, and any k prime to ϵ, every integer can be expressed in infinitely many ways as a sum of five values of the function*

$$p(x) = \epsilon(x^3 - x)/6 + kx + c$$

for integral values of x.

THEOREM 2. *For any k and c, every integer can be expressed as a sum of four values of the function*

$$P(x) = (x^3 - x)/6 + kx + c$$

for integral values of x.

Theorem 1 is trivially true when $\epsilon = 0$. For, in this case, $(\epsilon, k) = 1$ implies $k = 1$. Henceforth we take $\epsilon \neq 0$.

---

[1] Presented to the Society, April 27, 1940.

[2] *On the representation of integers by the sums of seven cubic functions*, Tôhoku Mathematical Journal, vol. 41 (1935–1936), pp. 361–366.

[3] Hua fails to mention in his formulation of this theorem whether his ϵ may take negative values. It seems that Hua implicitly assumed ϵ positive, as was noted by Pall in his review of the Hua paper in the Zentralblatt für Mathematik, vol. 14, p. 10. (This assumption was probably unnecessary, however.)

Let us first prove Theorem 1 for Hua's polynomial $f(x)$. Since x^3-x is divisible by 6, it is evident that

$$\frac{\epsilon}{6}(x^3-x)+x \equiv x \pmod{|\epsilon|}, \tag{1}$$

for any integer x. Hence it is evident that there exist $|\epsilon|$ incongruent values modulo $|\epsilon|$ of f. Consequently for any integer n it is true that there exist integers t and m such that

$$n = f(t) + m\epsilon. \tag{2}$$

We can show that $m\epsilon$ is a sum of four values of f by employing the following identity:

$$f(m+1)+f(m-1)+2f(-m) = m\epsilon. \tag{3}$$

From (2) and (3), we get

$$n = f(t)+f(m+1)+f(m-1)+2f(-m), \tag{4}$$

which tells us that every integer is a sum of five values of $f(x)$, x integral. Since we can replace t in (2) by any number congruent to it modulo $|\epsilon|$, it follows that the representation (4) can be made in infinitely many ways.

With but slight modification of this proof, we can obtain the more general result stated in Theorem 1. In this case (1) becomes $p(x) \equiv kx+c \pmod{|\epsilon|}$. And, since k is prime to ϵ, $kx+c$ ranges over a complete set of residues modulo $|\epsilon|$ when x does. Now if we replace f by p in (3), the right member of this identity becomes $m\epsilon+4c$. From these facts we can deduce Theorem 1.

If we replace f by P in (3), the right member of our identity becomes $m+4c$, which ranges over all of the integers with m; hence we have proved Theorem 2. It is evident from Theorem 1, with $\epsilon=1$, that we could duplicate Hua's "infinitely many ways" in our Theorem 2 if we increased the number of summands to five.

COOPER UNION INSTITUTE OF TECHNOLOGY

DIVISORS OF ZERO IN MATRIC RINGS

NEAL H. McCOY

1. **Introduction.** An element a of a ring S is a *divisor of zero* in S if there exists a nonzero element x of S such that $ax=0$, or a nonzero element y of S such that $ya=0$. The purpose of the present note is to obtain several theorems about divisors of zero in matric rings which, although quite elementary in character, have apparently not been previously noted. Throughout, unless otherwise stated, R will be used to denote an arbitrary commutative ring with unit element. Let R_n denote the ring of all matrices of order n with elements in R, and $R[\lambda]$ the ring of polynomials in the indeterminate λ with coefficients in R. If A is an element of R_n, we shall denote by $f(\lambda)=|\lambda-A|$ the *characteristic polynomial* of A, and thus $f(\lambda)$ is an element of $R[\lambda]$, with leading coefficient 1. The ideal $\mathfrak{m}$ of all elements $g(\lambda)$ such that $g(A)=0$ is the *minimum ideal* of A. If the minors of $\lambda-A$ of order $n-1$ are denoted by $h_{ij}(\lambda)$ $(i, j=1, 2, \cdots, n)$, it has been shown in a previous paper[1] that $g(\lambda)\equiv 0$ $(\mathfrak{m})$, if and only if

$$g(\lambda)h_{ij}(\lambda)\equiv 0 \quad (f(\lambda)), \qquad i, j = 1, 2, \cdots, n.$$

If $R[A]$ denotes the subring of R_n generated by A together with the unit element of R_n, which we identify with the unit element of R, then the elements of $R[A]$ are the polynomials in A with coefficients in R. It is quite easy to show that A is a divisor of zero in R_n if and only if $|A|$ is a divisor of zero in R. But we shall show, in §2, that A is actually a divisor of zero in $R[A]$ if it is a divisor of zero in R_n—a fact which is almost trivial if R is a field. This theorem is used in the following section in which we define, by means of the Sylvester determinant, the *resultant* $\mathfrak{R}(f, g)$ of two elements $f(\lambda)$ and $g(\lambda)$ of $R[\lambda]$ and show, following Frobenius, that if $f(\lambda)$ has leading coefficient unity,

$$\mathfrak{R}(f, g) = |g(A)|,$$

where A is *any* matrix having $f(\lambda)$ as characteristic polynomial. It then follows readily that an element $g(\lambda)$ of $R[\lambda]$ is *prime to*[2] $\mathfrak{m}$ if and

[1] Neal H. McCoy, *Concerning matrices with elements in a commutative ring*, this Bulletin, vol. 45 (1939), pp. 280–284. Hereafter, this paper will be referred to as M.

[2] For definitions of this term, see W. Krull, *Idealtheorie in Ringen ohne Endlichkeitsbedingung*, Mathematische Annalen, vol. 101 (1929), pp. 729–744. This will be referred to later as K. A definition will also be found in §3 of the present paper.

only if it is prime to $f(\lambda)$, which in turn is true if and only if $\mathfrak{R}(f,g)$ is not a divisor of zero in R.

If A is such that $\mathfrak{m}=(f(\lambda))$, we may say, following Sylvester, that A is *not derogatory*. Let us, as above, denote the minors of $\lambda-A$ of order $n-1$ by $h_{ij}(\lambda)$, and set

$$\mathfrak{a} = (h_{11}(\lambda), h_{12}(\lambda), \cdots, h_{nn}(\lambda)).$$

We conclude our remarks by showing that, if each ideal in R has a finite basis, A is not derogatory if and only if $\mathfrak{a}$ contains an element of R which is not a divisor of zero in R.

2. **Divisors of zero in $R[A]$.** We shall now prove

THEOREM 1. *An element A of R_n is a divisor of zero in $R[A]$ if and only if $|A|$ is a divisor of zero in R.*

The necessity follows easily by a familiar argument. Suppose that A is a divisor of zero in $R[A]$ and hence that there exists a nonzero element X of $R[A]$ such that $AX=0$. In other words, if $A=(a_{ij})$, the following system of equations,

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \qquad i = 1, 2, \cdots, n,$$

has a solution $(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)$ in R with some $\bar{x}_\alpha \neq 0$. Multiply these equations in order by the respective cofactors of $a_{1\alpha}, a_{2\alpha}, \cdots, a_{n\alpha}$ in $|A|$, and add. There results $\bar{x}_\alpha|A|=0$, and hence $|A|$ is a divisor of zero in R.

We now pass to the second part of the proof. Accordingly, we assume that there is an element $k\neq 0$ of R such that $k|A|=0$, and shall show that A is a divisor of zero in $R[A]$. Our method is to assume that A is not a divisor of zero in $R[A]$ and obtain a contradiction.

First, we prove the following

LEMMA. *If B and C are elements of R_n and k is an element of R such that $kB=kC$, then $k|B|=k|C|$.*

This follows from the fact that $kb_{ij}=kc_{ij}$, where $B=(b_{ij})$ and $C=(c_{ij})$, together with the following calculation. If $b=b_{i_11}b_{i_22}\cdots b_{i_nn}$ is any term, except possibly for sign, in the expansion of $|B|$, and $c=c_{i_11}c_{i_22}\cdots c_{i_nn}$ the corresponding term in the expansion of $|C|$, then

$$\begin{aligned} kb &= (kb_{i_11})b_{i_22}\cdots b_{i_nn} = (kc_{i_11})b_{i_22}\cdots b_{i_nn} \\ &= c_{i_11}(kb_{i_22})b_{i_33}\cdots b_{i_nn} = c_{i_11}c_{i_22}(kb_{i_33})\cdots b_{i_nn}, \end{aligned}$$

and a continuation of this process shows finally that this is equal to kc.

We now turn to the proof of Theorem 1. The matrix A satisfies its characteristic equation,[3] and hence there is a relation of the form

$$A^n + a_1A^{n-1} + \cdots + a_n = 0, \tag{1}$$

where $a_n = \pm |A|$, and thus $ka_n = 0$. Hence, from (1), we get an equation of the form

$$kA(A^{n-1} + a_1A^{n-2} + \cdots + a_{n-1}) = 0.$$

But, by our assumption that A is not a divisor of zero in $R[A]$, it follows that

$$k(A^{n-1} + a_1A^{n-2} + \cdots + a_{n-1}) = 0, \tag{2}$$

or, as we may write it,

$$kA(A^{n-2} + a_1A^{n-3} + \cdots + a_{n-2}) = -ka_{n-1}.$$

Now, applying the lemma, we see that

$$k|A| \cdot |A^{n-2} + a_1A^{n-3} + \cdots + a_{n-2}| = k(-1)^n a_{n-1}^n.$$

But $k|A| = 0$, and therefore $ka_{n-1}^n = 0$. Suppose $ka_{n-1}^j \neq 0$, but $ka_{n-1}^{j+1} = 0$ $(j \geqq 0)$. Then $k' = ka_{n-1}^j \neq 0$ has the property that $k'a_n = k'a_{n-1} = 0$, and hence from (2),

$$k'A(A^{n-2} + \cdots + a_{n-2}) = 0.$$

A repetition of this argument shows that there is an element $k'' \neq 0$ of R, which is a multiple of k', such that $k''a_n = k''a_{n-1} = k''a_{n-2} = 0$. Continuing this process we finally get an element $l \neq 0$ of R such that $lA = 0$, which violates our assumption that A is not a divisor of zero in $R[A]$. The theorem is therefore established.

Our method of proof also establishes the following

COROLLARY. *If A is a divisor of zero in R_n, there exists an element $g(A)$ of $R[A]$ such that*

$$Ag(A) = 0, \qquad g(A) \neq 0,$$

where $g(\lambda)$ is an element of $R[\lambda]$ with the property that, for some positive integer i,

$$\lambda^i g(\lambda) = kf(\lambda),$$

$f(\lambda)$ being the characteristic polynomial of A and k an element of R.

[3] See M, p. 282.

3. **Resultants and their properties.** We first make some preliminary remarks.[4] Let $\mathfrak{a}$ be an ideal in the commutative ring R, with unit element. An element b of R is said to be *prime to* $\mathfrak{a}$ if $bx \equiv 0$ $(\mathfrak{a})$ implies that $x \equiv 0$ $(\mathfrak{a})$, otherwise b is *not prime to* $\mathfrak{a}$. An element b is prime to the element a if b is prime to the principal ideal (a). If every element of an ideal $\mathfrak{b}$ is not prime to $\mathfrak{a}$, but every proper divisor of $\mathfrak{b}$ contains an element prime to $\mathfrak{a}$, then $\mathfrak{b}$ is necessarily a prime ideal which we shall call a *maximal prime ideal belonging to* $\mathfrak{a}$. Krull has shown, by methods which are necessarily transfinite if R is unrestricted, that every ideal has maximal prime ideal divisors and that *the elements of R which are not prime to* $\mathfrak{a}$ *are precisely the elements which are contained in some maximal prime ideal belonging to*[5] $\mathfrak{a}$.

We now consider the ring $R[\lambda]$, where λ is an indeterminate, and let

$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n, \qquad g(\lambda) = b_0\lambda^m + b_1\lambda^{m-1} + \cdots + b_m$$

be two elements of $R[\lambda]$, the leading coefficient of $f(\lambda)$ being 1. In case R is a field, the vanishing of the resultant of $f(\lambda)$ and $g(\lambda)$ furnishes a necessary and sufficient condition that $f(\lambda)$ and $g(\lambda)$ be not prime to each other.[6] We shall now obtain an analogous result in case R is unrestricted.

Let us denote the resultant of $f(\lambda)$ and $g(\lambda)$ by $\mathfrak{R}(f, g)$ and define it in a purely formal way by the Sylvester determinant

$$\mathfrak{R}(f, g) = \begin{vmatrix} 1 & a_1 & \cdot & \cdot & \cdot & a_n & & \\ & 1 & a_1 & \cdot & \cdot & \cdot & a_n & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & 1 & a_1 & \cdot & \cdot & \cdot & a_n \\ b_0 & b_1 & \cdot & \cdot & \cdot & b_m & & \\ & b_0 & b_1 & \cdot & \cdot & \cdot & b_m & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & b_0 & b_1 & \cdot & \cdot & \cdot & b_m \end{vmatrix},$$

the blank spaces consisting of zeros.

The connection with the preceding part of the paper is indicated by

[4] Cf. K, p. 732.

[5] In fact, it can be shown that if in R every ideal has a finite basis, the maximal prime ideals belonging to a are precisely the maximal among the prime ideals associated with the primary ideals in a normalized decomposition of $\mathfrak{a}$ into the intersection of primary ideals. All of these prime ideals are the *prime ideals belonging to* $\mathfrak{a}$.

[6] In case R is a field, this implies that $f(\lambda)$ and $g(\lambda)$ have a common factor of degree at least one.

THEOREM 2. *If A is any element of R_n which has characteristic polynomial $f(\lambda)$, then*

$$\mathfrak{R}(f, g) = |g(A)|. \tag{3}$$

This is an almost immediate consequence of Frobenius' theorem for the case in which R is the field of complex numbers.[7] For if we consider the elements of a matrix A, as well as the coefficients of $g(\lambda)$ to be independent complex variables, relation (3) is seen to be an identity in these variables and thus remains true if these variables are replaced by elements of[8] R.

Now let A be any fixed element of R_n with characteristic polynomial $f(\lambda)$. The ideal $\mathfrak{m}$ of all elements $h(\lambda)$ of $R[\lambda]$ such that $h(A)=0$ is the minimum ideal of A. It is now clear that an element $g(\lambda)$ of $R[\lambda]$ is prime to $\mathfrak{m}$ if and only if $g(A)$ is not a divisor of zero in $R[A]$, and Theorem 1 states that this is the case if and only if $|g(A)|$ is not a divisor of zero in R. We shall now prove

THEOREM 3. *An element $g(\lambda)$ of $R[\lambda]$ is prime to $\mathfrak{m}$ if and only if it is prime to $f(\lambda)$.*

By the preceding remarks, and Theorem 2, we see that $g(\lambda)$ is prime to $\mathfrak{m}$ if and only if $\mathfrak{R}(f, g)$ is not a divisor of zero in R. But clearly $\mathfrak{R}(f, g)$ depends only on $f(\lambda)$ and $g(\lambda)$ and not on the particular choice of matrix A with characteristic polynomial $f(\lambda)$. Let A' be a matrix with characteristic polynomial $f(\lambda)$, and such that its minimum ideal is[9] $\mathfrak{m}'=(f(\lambda))$. Then $g(\lambda)$ is prime to $\mathfrak{m}'$ if and only if $\mathfrak{R}(f, g)$ is not a divisor of zero in R. But it was found above that this is precisely the condition that $g(\lambda)$ be prime to $\mathfrak{m}$, and the theorem is established.

Since $g(\lambda)$ is prime to $\mathfrak{m}$ if and only if $\mathfrak{R}(f, g)$ is not a divisor of zero in R, we have incidentally proved the following result which is independent of the theory of matrices:

THEOREM 4. *If R is a commutative ring with unit element and λ is an indeterminate, an element $g(\lambda)$ of $R[\lambda]$ is prime to the element $f(\lambda)$, with leading coefficient unity, if and only if $\mathfrak{R}(f, g)$ is not a divisor of zero in R.*

4. **A characterization of matrices which are not derogatory.** Let

[7] G. Frobenius, *Ueber lineare Substitutionen und bilineare Formen*, Journal für die reine und angewandte Mathematik, vol. 84 (1878), p. 11.

[8] Cf. M, p. 281.

[9] This will certainly be the case if A' is the companion matrix of $f(\lambda)$. See relation (4) of §4.

A be an element of R_n, with characteristic polynomial $f(\lambda)$ and minimum ideal $\mathfrak{m}$. Following Sylvester, we may say that A is *not derogatory* if $\mathfrak{m}=(f(\lambda))$. Let $h_{ij}(\lambda)$ denote the minors of $\lambda-A$ of order $n-1$. It was remarked above that an element $g(\lambda)$ of $R[\lambda]$ is an element of $\mathfrak{m}$ if and only if

$$g(\lambda)h_{ij}(\lambda) \equiv 0 \quad (f(\lambda)), \qquad i, j = 1, 2, \cdots, n. \tag{4}$$

We shall now make use of this fact in a proof of the following theorem:

THEOREM 5. *If in R each ideal has a finite basis, the matrix A is not derogatory if and only if the ideal*

$$\mathfrak{a} = (h_{11}(\lambda), h_{12}(\lambda), \cdots, h_{nn}(\lambda))$$

contains an element of R which is not a divisor of zero.

Suppose b is an element of R which is not a divisor of zero and that $b\equiv 0$ $(\mathfrak{a})$. Then, from (4), it follows that if $g(\lambda)\equiv 0$ $(\mathfrak{m})$, then $bg(\lambda)\equiv 0$ $(f(\lambda))$. But $f(\lambda)$ has leading coefficient 1, and from this fact, and the fact that b is not a divisor of zero, an easy calculation shows that $g(\lambda)\equiv 0$ $(f(\lambda))$. Since always $f(\lambda)\equiv 0$ $(\mathfrak{m})$, this shows that A is non-derogatory.

Now let us assume that A is not derogatory which means that $g(\lambda)\mathfrak{a}\equiv 0$ $(f(\lambda))$ implies that $g(\lambda)\equiv 0$ $(f(\lambda))$. Under hypothesis of a finite basis for each ideal in R, this means that $\mathfrak{a}$ is not divisible by any prime ideal belonging to[10] $(f(\lambda))$. This implies, in turn, that there is an element $q(\lambda)$ of $\mathfrak{a}$ which is not in any prime ideal belonging to $(f(\lambda))$ and is thus prime to[11] $f(\lambda)$. But since $q(\lambda)$ is prime to $f(\lambda)$, it follows by Theorem 4 that $\mathfrak{R}(f, q)=|q(A)|$ is not a divisor of zero in R. Let $|q(A)|=a$. Extend R to a ring R' consisting of elements of the form r/a^i $(i=0, 1, \cdots; r \text{ in } R)$.[12] In R', a has an inverse and therefore $q(A)$ has an inverse[13] which is a polynomial in A with coefficients in R', say

$$p(A) = (b_0/a^{i_0})A^k + (b_1/a^{i_1})A^{k-1} + \cdots + (b_k/a^{i_k}).$$

Since $q(A)p(A)=1$, in R we have a relation of the form

$$q(A)(c_0A^k + c_1A^{k-1} + \cdots + c_k) = a^l = b,$$

[10] Cf. van der Waerden, *Moderne Algebra*, vol. 2, p. 41.

[11] In fact, it may be shown by induction on k that if every element of an ideal $\mathfrak{a}$ is contained in some one of the prime ideals $\mathfrak{p}_1, \mathfrak{p}_2, \cdots, \mathfrak{p}_k$, then $\mathfrak{a}$ is divisible by some one of the $\mathfrak{p}_i$.

[12] See R. Holzer, *Zur Theorie der primären Ringe*, Mathematische Annalen, vol. 96 (1927), p. 722.

[13] Cf. M, p. 282.

where b is not a divisor of zero. If we set

$$t(\lambda) = c_0\lambda^k + c_1\lambda^{k-1} + \cdots + c_k,$$

we have in $R[\lambda]$

$$q(\lambda)t(\lambda) \equiv b \quad (\mathfrak{m}),$$

which, by hypothesis, implies that

$$q(\lambda)t(\lambda) \equiv b \quad (f(\lambda)).$$

Now $q(\lambda) \equiv 0$ $(\mathfrak{a})$ and clearly also $f(\lambda) \equiv 0$ $(\mathfrak{a})$, from which it follows that $b \equiv 0$ $(\mathfrak{a})$. The theorem is therefore established.

SMITH COLLEGE